U.PORTO
INSTITUTO DE CIÊNCIAS BIOMÉDICAS ABEL SALAZAR
UNIVERSIDADE DO PORTO

MOLECULAR PATHOPHYSIOLOGY UNDERLYING NEURONAL CEROID LIPOFUSCINOSES: CLN2 AND CLN5

CARLOS JORGE PEREIRA BESSA

Dissertação de doutoramento em Ciências Biomédicas

2009

Carlos Jorge Pereira Bessa

Molecular Pathophysiology Underlying Neuronal Ceroid Lipofuscinoses: CLN2 and CLN5

Dissertação de Candidatura ao grau de Doutor em Ciências Biomédicas submetida ao Instituto de Ciências Biomédicas de Abel Salazar da Universidade do Porto.

Orientador – Doutora Maria Gil Roseira Ribeiro
Categoria – Técnica Superior de Saúde
Afiliação – Instituto Nacional de Saúde Dr. Ricardo Jorge, Centro de Genética Médica Jacinto Magalhães

Co-orientador – Doutor Jorge Eduardo da Silva Azevedo
Categoria – Professor Catedrático
Afiliação – Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto

Legal precepts

Preceitos legais

In accord with the 2nd of the 8th article of Decreto-lei nº 388/70, the experimental results stated bellow in Publications I/II/III were used in this dissertation. The author of this dissertation declares his participation in the outlining and execution of the experimental work, in the data presentation and preparation of the work referenced below under the name **Bessa, C** / **Carlos Bessa**:

De acordo com o disposto no nº2 do artigo 8º do Decreto-lei nº388/70, nesta dissertação foram utilizados os resultados dos trabalhos publicados abaixo referenciados. No cumprimento do disposto no referido Decreto-lei, o autor desta dissertação declara que interveio na concepção e execução do trabalho experimental, na interpretação e redacção dos resultados publicados sob o nome **Bessa, C / Carlos Bessa**:

Publication I / Publicação I

Bessa C, Teixeira CA, Dias A, Alves M, Rocha S, Lacerda L, Loureiro L, Guimarães A, Ribeiro MG. *CLN2/TPP1 deficiency: the novel mutation IVS7-10A>G causes intron retention and is associated with a mild disease phenotype.* Mol Genet Metab. **2008** Jan;93(1):66-73.

Publication II / Publicação II

Bessa C*, Teixeira CA*, Mangas M, Dias A, Sá Miranda MC, Guimarães A, Ferreira JC, Canas N, Cabral P, Ribeiro MG. *Two novel CLN5 mutations in a Portuguese patient with vLINCL: insights into molecular mechanisms of CLN5 deficiency.* Mol Genet Metab. **2006** Nov;89(3):245-53. * These authors contributed equally to the work.

Publication III / Publicação III

Mia-Lisa Schmiedt*, **Carlos Bessa***, Claudia Heine, Maria Gil Ribeiro, Anu Jalanko, Aija Kyttälä. *The CLN5 protein: New insights into cellular processing and transport.* **2009***. Submitted. * These authors contributed equally to the work.*

This work received financial support from Fundação para a Ciência e Tecnologia (FCT), through a PhD fellowship (SFRH/BD/17560/2004), and from Vaincre les Maladies Lysosomales society.

Este trabalho foi financiado pela Fundação para a Ciência e Tecnologia (FCT), através de uma bolsa de doutoramento (SFRH/BD/17560/2004), e pela associação Vaincre les Maladies Lysosomales.

To my family

Table of Contents

1 Abbreviations

aa	amino acid(s)
ANCL	adult neuronal ceroid lipofuscinosis
BFA	brefeldin A
bp	base pair
BSA	bovine serum albumin
CD-MPR	cation dependent mannose-6-phosphate receptor
cDNA	complementary deoxyribonucleic acid
CHX	cycloheximide
CI-MPR	cation independent mannose-6-phosphate receptor
CLN	ceroid lipofuscinosis, neuronal (gene nomenclature)
CLN	ceroid lipofuscinosis, neuronal (protein nomenclature)
CLN5	human ceroid lipofuscinosis, neuronal 5 protein
Cln5	mouse ceroid lipofuscinosis, neuronal 5 protein
CNS	central nervous system
COL	R112H mutation in the *CLN5* gene (Colombian)
COS-1 cells	African green monkey kidney cells
CTSD	cathepsin D
DNA	deoxyribonucleic acid
EPMR	progressive epilepsy with mental retardation
ER	endoplasmic reticulum
ERT	enzyme replacement therapy
ESE	exonic splicing enhancer
EUR	p.D279N mutation in *CLN5* gene (European)
Fin_M	p.Y392X mutation in *CLN5* gene (Finnish major)
GROD	granular osmophilic deposits

GalCer	galactosylceramide
HeLa cells	human cervical tumor cells
INCL	infantile neuronal ceroid lipofuscinosis
ISE	intronic splicing enhancer
JNCL	juvenile neuronal ceroid lipofuscinosis
kb	kilobase
kDa	kilodalton
LAMP	lysosomal associated membrane protein
LGP	lysosomal glycoprotein
LIMP	lysosomal integral membrane protein
LINCL	late infantile neuronal ceroid lipofuscinosis
LMP	lysosomal membrane protein
LSD	lysosomal storage disease
M6P	mannose-6-phosphate
MHC	major histocompatibility complex
MPR	mannose-6-phosphate receptor
MPR -/-	mannose-6-phosphate receptor deficient
MFS	major facilitator superfamily of proteins
mRNA	messenger RNA
MW	molecular weight
NCL	neuronal ceroid lipofuscinosis
NSF	N-ethylmaleimide sensitive factor
nt	nucleotide
NTK cells	natural killer T cells
PCR	polymerase chain reaction
PDI	protein disulfide isomerase
POR	p.R112P mutation in *CLN5* gene (Portuguese)
PPT1	palmitoyl protein thioesterase 1 (CLN1)

RNA	ribonucleic acid
SAP	sphingolipid activator protein (saposin)
SCMAS	sub-unit C of mitochondrial ATPase
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SH-SY5Y cells	human neuroblastoma cells
SNAP	soluble N-ethylmaleimide sensitive factor attachment protein
ss	splice site
SSCP	single strand conformational polymorphisms
SWE	p.E253X mutation in *CLN5* gene (Swedish)
TCA	tri-carboxilic acid
TGN	*trans*-golgi network
TPP1	tri-peptidyl peptidase 1 (CLN2)
vLINCL	variant late infantile neuronal ceroid lipofuscinosis
$vLINCL_{Fin}$	Finnish variant late infantile neuronal ceroid lipofuscinosis
wt	wild-type

2 Abstract

The neuronal ceroid lipofuscinoses (NCLs) are the most common group of neurodegenerative disorders in childhood. The incidence has been estimated at 1:65000 live births in Portugal, and between 1:12500 and 1:100000 worldwide. The disease hallmark is the lysosomal accumulation of an autofluorescent pigment, similar to ceroid-lipofuscin, in several cell types and tissues. Clinical features include progressive blindness, epilepsy, and mental and motor impairment leading to premature death. The first clinical manifestations may occur soon after birth or until adulthood. NCL diseases with characterised gene defects are recessively inherited and to date, eight genes are known. Mutations in *CLN2*, *CLN5*, *CLN6* and *CLN7* genes have been reported in patients with the late infantile phenotype. Despite the genetic heterogeneity, clinical and neuropathological findings are highly similar and suggest common molecular mechanisms behind NCL disorders. Although the first gene defects were described more than ten years ago, the molecular pathological mechanisms underlying NCL have not yet been completely elucidated.

The LINCL is associated with mutations in the *CLN2* gene, which codes for the soluble lysosomal enzyme tripeptidyl peptidase 1 (TPP1), known to remove peptides from the N-terminus of small polypeptides. However, the *in vivo* substrates are still unknown. Mutations in the *CLN5* gene have been associated with the Finnish variant of the late infantile form ($vLINCL_{Fin}$) as the disease is particularly prevalent in the Finnish population. The CLN5 protein has been described as a soluble lysosomal glycoprotein but its cell biology is largely unknown.

The aim of this thesis was to contribute to elucidate the molecular pathophysiology of the CLN2 and CLN5 deficiency observed in patients with the classical and the Finnish variant LINCL, respectively. Furthermore, the work focused on the characterization of some properties of the CLN5 protein necessary for the study of the molecular and cellular mechanisms of $vLINCL_{Fin}$ causing mutations.

The large majority of mutations reported in the *CLN2* gene are associated with the classical LINCL. In this study a novel splice site mutation in the *CLN2* gene, IVS7-10A>G, was identified in a patient exhibiting an atypical later LINCL phenotype. The mutation was demonstrated to cause an aberrant splicing of exon 8 by creating a novel acceptor splice site. However no detectable impact on the level of *CLN2*/mRNA or protein maturation was noticed, although a slight reduction in the protein level was observed. The mRNA specie with the intron 7 sequence retention is predicted to code for a protein with three extra

amino acids between Pro295 and Gly296, near the catalytic triad of TPP1. In fact, the mutation dramatically affected the *in vitro* enzymatic activity. Nevertheless, the mutant TPP1 of either leucocytes or fibroblasts extracts displayed an activity approximately 2-fold higher when compared to classical LINCL patient's cells carrying the p.R208X mutation. The retention of a similar residual enzymatic activity *in vivo* dictates the molecular basis for the late onset and delayed progression of the disease observed in the patient.

Novel disease-causing mutations were also identified in the *CLN5* gene: c.335G>C (p.R112P) and c.565C>T (p.Q189X). Those alterations were found in a Portuguese patient, thus representing the first evidence of the disease in this country and the third description in non northern European countries. At the time the number of known mutations in the *CLN5* gene increased from five to seven. The existence of a non-producing allele carrying the p.Q189X mutation supported the reduced the level of *CLN5*/mRNA observed in patient's fibroblasts. With the purpose to further investigate the impact of mutations on the molecular and cellular properties of CLN5, cell over expressing systems were generated and the protein studied with a novel peptide-specific antibody raised against the C-terminal segment of the human protein. The CLN5 protein was found to be N-terminally processed, N-glycosylated with high-mannose and complex type oligosaccharides, and targeted to lysosomes. All analyzed disease mutations (p.R112H, p.R112P, p.E253X, p.D297N and p.Y392X) disrupted the lysosomal trafficking. Interestingly, mutant polypeptides were observed in the ER and found to be sensitive to EndoH and PGNaseF digestion, although for some mutations trafficking to lysosomes was allowed in some extent. However, the degree of lysosomal targeting did not correlate to the disease onset. PPT1 over expression was found to facilitate trafficking of mutant CLN5 polypeptides to lysosomes. Overall, these data suggest that a retrograde transport event from the Golgi to the ER may be necessary for correct CLN5 processing and maturation, raising the hypothesis of CLN5 may also function outside lysosomes.

Overall, the data described in this thesis made two contributions in this particular field of the scientific knowledge: detailed analysis on the biological consequences of novel *CLN2* and *CLN5* mutations provided a better understanding of the molecular and cellular pathological processes behind the disease; data on the biological properties of the CLN5 protein raised novel questions with implication for the whole NCL proteome. The characterization of the cell biology of NCL proteins, including the identification of their interacting partners, are issues that will continue to be investigated in the future as this information will be crucial to enlighten the NCL disease mechanisms.

Keywords: Neuronal ceroid lipofuscinoses; Late infantile neuronal ceroid lipofuscinosis; Finnish variant LINCL; CLN2; Splicing mutations; TPP1 deficiency; CLN5; Lysosomal trafficking

3 Resumo

A Ceroido-lipofuscinose neuronal (CLN) é o grupo de doenças neurodegenerativas mais frequente na infância. Em Portugal a incidência foi estimada em 1:65000 nados vivos, variando a nível mundial entre 1:12500 e 1:100000. A CLN caracteriza-se pela acumulação intracelular, em diversos tecidos, de um material autofluorescente com características físico-químicas semelhantes às da lipofuscina. Clinicamente esta doença é caracterizada por cegueira progressiva, epilepsia, disfunção motora e cognitiva, observando-se morte prematura. A idade de aparecimento dos primeiros sinais clínicos varia entre o período pós-natal e a idade adulta. Actualmente são conhecidas oito variantes genéticas, apresentando um padrão de transmissão autossómico recessivo. O fenótipo infantil tardio (LINCL) é geneticamente heterogéneo, encontrando-se associado a mutações em quatro genes: *CLN2*, *CLN5*, *CLN6* e *CLN7*. Apesar da heterogeneidade genética, as observações clínicas e neuropatológicas sugerem um mecanismo molecular comum subjacente a este grupo de doenças. Embora as primeiras descrições da sua base molecular genética tenham sido publicadas há mais de 10 anos, os mecanismos fisiopatológicos da CLN ainda não estão completamente esclarecidos.

A forma clássica LINCL está associada a mutações no gene *CLN2*. Este gene codifica para a proteína lisossomal solúvel tripeptidil-peptidase 1 (TPP1) que remove péptidos da extremidade N-terminal de proteínas. Contudo, os seus substratos fisiológicos não são ainda conhecidos. Mutações no gene *CLN5* estão associadas à variante Finlandesa da forma clássica LINCL ($vLINCL_{Fin}$), que é particularmente frequente na população Finlandesa. A função da proteína CLN5, descrita como uma glicoproteína lisossomal solúvel, permanece ainda desconhecida.

Este projecto foi desenvolvido com o objectivo de contribuir para o conhecimento dos mecanismos moleculares fisiopatológicos da deficiência em CLN2 e CLN5, associadas, respectivamente, à forma clássica e à variante Finlandesa de LINCL. Procurou-se, ainda, caracterizar algumas das propriedades da proteína CLN5 essenciais para a compreensão dos mecanismos celulares e moleculares de mutações causais no gene *CLN5*.

A maioria das mutações descritas no gene *CLN2* estão associadas à forma LINCL. Neste estudo, foi identificada uma nova mutação de *splicing* no gene *CLN2*, IVS7-10A>G, observada num doente com um fenótipo tardio, atípico, de LINCL. Esta mutação origina o aparecimento de um possível local aceitador de *splicing*. A abordagem

experimental demonstrou que esse local é usado *in vivo*, resultando no processamento anormal do exão 8. Não foi detectado qualquer impacto no nível do ARNm/*CLN2* ou da maturação da proteína, apesar de se ter observado uma ligeira redução do nível da proteína mutada. O ARNm mutado codifica para uma proteína com 3 aminoácidos adicionais entre o aminoácido Pro295 e Gly296, que se localizam próximo da tríade catalítica de TPP1. De facto, em extractos celulares de leucócitos e de fibroblastos foi observada uma deficiente actividade enzimática *in vitro* da proteína mutada. No entanto, esse nível residual de actividade era cerca de duas vezes superior ao observado em extractos celulares de doentes com a forma clássica infantil-tardia e homozigóticos para a mutação p.R208X. Um nível de actividade enzimática *in vivo* semelhante ao observado *in vitro* constituirá a base molecular para o fenótipo tardio e progressivo observado no doente analisado.

Neste estudo foram também identificadas duas novas mutações causais no gene *CLN5*: c.335G>C (p.R112P) e c.565C>T (p.Q189X). Essas alterações foram identificadas pela primeira vez num doente Português e representaram a terceira descrição de $vLINCL_{Fin}$ em áreas geográficas onde a doença não é prevalente, aumentando o número de mutações descritas no gene de cinco para sete. A existência de um alelo nulo, com a mutação p.Q189X, explica o nível reduzido de ARNm/*CLN5* observado em fibroblastos do doente. No sentido de compreender o impacto dessas mutações *missense* na biologia celular da proteína CLN5 foram criados sistemas celulares de sobre expressão e a proteína foi estudada com um novo anticorpo específico para a região C-terminal da proteína humana. Os resultados obtidos revelaram que a proteína CLN5 sofre um processamento N-terminal, correspondendo provavelmente a remoção do péptido sinal no RE, sendo transportada para o sistema endossomal/lisossomal. Todas as mutações estudadas (p.R112H, p.R112P, p.E253X, p.D297N e p.Y392X) resultaram na interrupção do tráfego lisossomal. Os polipéptidos mutados, sensíveis à digestão por EndoH e PGNaseF, foram observados principalmente no ER, e em alguns casos, também nos endossomas/lisossomas. No entanto, não foi possível correlacionar o nível de proteína observada no compartimento lisossomal e o fenótipo clínico associado a cada mutação. De referir, ainda, que a sobre expressão da proteína PPT1 permitiu a correcção do tráfego lisossomal dos polipéptidos mutados. Estes resultados sugerem que o transporte da proteína do aparelho de Golgi para o ER poderá ser um requisito importante para o processamento e maturação celular da proteína CLN5, não sendo, presentemente, de excluir a possibilidade desta proteína ser funcionalmente relevante em compartimentos extra-lisossomais.

Os resultados descritos nesta tese proporcionaram duas contribuições nesta área

particular do conhecimento científico: o estudo detalhado das consequências moleculares de mutações nos genes *CLN2* e *CLN5* contribuiu para uma melhor compreensão dos processos moleculares e celulares subjacentes à doença; a caracterização das propriedades biológicas da proteína CLN5 colocou novas questões com implicações para todo o proteoma CLN. A caracterização da biologia celular das proteínas CLN e a identificação dos seus "parceiros" moleculares são linhas de estudo que continuarão a ser seguidas no futuro, sendo essenciais para a clarificação dos mecanismos patológicos por detrás da doença e para o desenvolvimento de novas terapias.

Palavras chave: Ceroido lipofuscinose neuronal; Forma infantil tardia (LINCL); Variante Finlandesa da forma LINCL; CLN2; Mutações de *splicing*; deficiência em TPP1; CLN5; Endereçamento lisossomal

4 Sommaire

Les lipofuscinoses neuronales ceroïdes (NCLs) sont le groupe le plus commun de maladies neurodégénératives dans l'enfance. L'incidence estimée est de 1,55 pour 100000 naissances vivantes au Portugal et entre 1:12500 et 1:100000 dans le monde. Les charactéritiques cliniques inclues cécité progressive, épilepsie, atrophie du cerveau et la dégénérescence moteure qui mène à la mort prémature. La caractéristique principale de la NCL est l'accumulation lysosomale d'un pigment autofluorescent, semblable à la lipofuscine céroïde, dans plusieurs types cellulaires et tissus. Les NCL avec anomalies génétiques caractérisées sont héréditaires récessives et à ce jour, huit gènes sont liés à ce groupe de maladies. Des mutations dans les gènes CLN2, CLN5, CLN6 et CLN7 ont été rapportés chez des patients avec le phénotype infantile tardive. Malgré l'hétérogénéité génétique, les résultats cliniques et neuropathologiques sont très semblables et suggèrent des mécanismes moléculaires communs à l'origine des NCLs. Bien que les premiers défauts génétiques ont été décris il y a dix ans, les mécanismes moléculaires sous-jacents à la pathologie des NCLs n'ont pas encore été complètement élucidés.

La forme infantile tardive classique des NCLs (LINCL) est associée à des mutations au gène CLN2 et représente environ 10% des cas portugais de NCL. CLN2 code pour l'enzyme soluble lysosomale tripeptidyl 1 peptidase (TPP1), qui permet éliminer des peptides de l'extrémité N-terminale de petits polypeptides. Cependant, in vivo, ses subtracts sont encore inconnus. Des mutations au gène CLN5 résultent dans la variante finlandaise de la forme infantile tardive (vLINCLFin) étant particulièrement réprésentée dans la population finlandaise. La protéine CLN5 a été décrite comme une glycoprotéine lysosomale mais la biologie cellulaire de cette protéine est largement inconnue.

Cette thèse vise contribuer à élucider la physiopathologie moléculaire de la carence de CLN2 et CLN5. En outre, les travaux ont abordé la caractérisation de certaines propriétés de la protéine CLN5 nécessaires pour l'étude des mécanismes moléculaires et cellulaires des mutations provoquant la vLINCLFin.

La grande majorité des mutations dans le gène CLN2 sont associées à la forme classique infantile tardive de la maladie. Dans cette étude, une nouvelle mutation dans un locale d'épissage dans le gène CLN2, IVS7-10A>G, a été identifiée chez un patient présentant un phénotype atypique plus tardive de LINCL. La mutation provoque un épissage aberrant de l'exon 8 en créant un site accepteur d'épissage nouveau. La mutation n'a eu aucun impact détectable sur le niveau de CLN2/mRNA et n'a pas d'incidence sur la maturation de la protéine, même si une légère diminution du niveau de

la protéine a été observée. L'ARNm avec la séquence de l'intron 7 est prévu coder une protéine avec trois acides aminés supplémentaires entre la proline 295 et la glycine 296, près de la triade catalytique de TPP1. En fait, la mutation affecte dramatiquement in vitro l'activité enzymatique de TPP1. Néanmoins, la protéine mutante a 'une activité résiduelle environ 2 fois plus élevée par rapport à un patient LINCL classique porteur de la mutation p.R208X, soit dans les leucocytes ou les fibroblastes. Ainsi, une activité similaire in vivo fournirait la base moléculaire de l'arrivée tardive et de la progression retardée de la maladie observées chez le patient.

Novelles mutations pathogéniques ont été également identifiées dans le gène CLN5: c.335G>C (p.R112P) et c.565C>T (p.Q189X). Ces altérations ont été observées chez un patient portugais représentant la troisième description de la maladie hors des pays d'Europe du Nord et la première preuve de la maladie dans la population portugaise, élevant au moment le nombre de mutations connues dans le gène CLN5 de cinq à sept. L'existence d'un allèle non producteur portant la mutation p.Q189X a été soutenue par la réduction du niveau de CLN5/mRNA observé dans les cellules fibroblastiques du patient. Avec le but d'approfondir l'impact de mutations faux-sens sur les propriétés moléculaires et cellulaires de la CLN5, des systèmes pour surexprimer la protéine ont été générés et étudiés avec un nouveau anticorps peptide-spécifique. La protéine CLN5 été procésée au N-terminale et envoyée pour le compartiment lysosomal. Toutes les mutations pathologiques analysées (p.R112H, p.R112P, p.E253X, p.D297N et p.Y392X) perturbaient le trafique lysosomal des polypeptides. Le degré de trafique lysosomale n'a pas, cependant, une corrélation avec l'apparition de la maladie suggérant ainsi que CLN5 peut également fonctioner dehors des lysosomes.

Dans l'ensemble, cette thèse a fait deux importantes contributions au domaine des NCLs: l'analyse détaillée des conséquences biologiques de nouvelles mutations à CLN2 et CLN5 permet une meilleure compréhension des processus moléculaires et cellulaires pathologiques derrière la maladie; les propriétés biologiques de la protéine CLN5 soulève des questions nouvelles ayant des répercussions sur l'ensemble du protéome NCL. La caractérisation de la biologie cellulaire des protéines NCL, y compris l'identification de leurs partenaires d'interaction, sont des questions qui continueront à être étudiés à l'avenir car cette information est crucial pour éclairer les mécanismes de la maladie NCL et favoriser le développement de thérapies futures.

Mots-clés: lipofuscinose céroïde; lipofuscinose ceroïde neuronale infantile tardive; LINCL variante finlandaise; CLN2; mutations d'épissage; carence en TPP1; CLN5; trafique lysosomale

5 Introduction

"Life is a complex thing"
Keyes 1966

During millions of years Life has evolved from "simple" single cell organisms to the complex living beings we can witness today. One cell dividing itself into two cells, and so on and so forth, resulting in the perceived diversity of cell types and singularities one can find in complex organisms, such as ourselves. At a particular time point in each cell, a myriad of biological molecules participate in numerous biochemical reactions required for it's development and survival. Within the beautiful complexity of the inner workings of cells an error in a single piece of the cellular machinery may result in grave repercussions that ultimately may lead to a process of disease to the organism.

"Cell genetics led us to investigate cell mechanics. Cell mechanics now compels us to infer the structures underlying it. (…) The theory of the cell revealed the unity of living processes; The study of the cell is beginning to reveal their physical foundations."
Darlington 1937

Medical science focus on the study of human diseases, applying that knowledge to prevent and cure maladies, thus improving our well being. But consequently, shedding light into pathological events helps to clarify fundamental processes in cell biology, enhancing our understanding of how humans and other animals function. In an age of widespread advances in health care and curative medicine that allow for increasingly longer lifespans, genetic, and more specifically neurological disorders, have assumed a great importance. Genetic neurodegenerative disorders are usually progressive and incurable, and the cellular events leading to neurodegeneration are still mostly unknown.

Neuronal Ceroid Lipofuscinosis (NCLs) are a group of rare, monogenic, recessively inherited neurodegenerative disorders affecting mostly children. The first clinical descriptions date back to the 19th century and the beginning of the 20th century (Haltia, 2006). The main characteristic feature is the accumulation of an autofluorescent storage material in many cell types, including neurons. Severe neuronal death comes associated with a clinical picture that includes epileptic seizures, progressive psicomotor deterioration, and visual failure leading to premature death (Siintola *et al*., 2006a). The

incidence has been estimated to range from 1:12.500 to 1:100.000 worldwide and at 1:65000 in Portugal (Haltia, 2003; Teixeira *et al*., 2003). To date 8 genes have been characterized as being associated with the disease and related to classical or variant forms of NCL based on the age of onset and clinical progression. NCL genes codify for diverse proteins presenting soluble or transmembrane topologies and various cellular distributions. The precise function of all NCL proteins is unknown, although they are speculated to participate in a common, yet unknown, biological pathway.

The work described in this thesis focus on the classical late infantile and the Finnish variant of the late infantile NCL form. The data describe contributions at an epidemiological level for the *CLN2* and *CLN5* genes, underlying LINCL and vLINCL$_{Fin}$ respectively, along with analysis on mutational mechanisms. Moreover, the study of the CLN5 protein brings novel insights on its biological properties, necessary for a better understanding of the impact of disease mutations and for the development of future disease therapies.

6 Review of the literature

6.1 Lysosomes

6.1.1 Structure and function of lysosomes

De Duve introduced the term lysosome in 1955, as a designation for litic particle or digestive body (de Duve *et al*., 1955). Lysosomes are distinguished from endosomes by the lack of mannose 6-phosphate receptors (MPRs) and cell surface recycling receptors. They are morphologically heterogeneous and appear in mammalian cells as organelles of ~ 0.5 µm in diameter representing 0.5-5 % of the total cell volume (Matteoni and Kreis, 1987). Proteomic analysis suggest the existence of at least 50-60 soluble hydrolases (Journet *et al*., 2002), responsible for the degradation of the numerous substrates that arrive at the lysosome from endocytosis, autophagy and even secretory material destined for destruction. The low pH associated with the lysosomal lumen (pH 4.6-5) is maintained by a conserved vacuolar membrane proton pump – V-type H+ - ATPase (Mellman *et al*., 1986).

The lysosomal membrane has a unique constitution, with a characteristic phospholipid composition (Bleistein *et al*., 1980) and is extremely rich in carbohydrates (Lloyd and Forster, 1986). Lysosomal associated membrane proteins (LAMPs) and lysosomal integral membrane proteins (LIMPs) are heavily glycosylated and are estimated to represent approximately 50 % of the total membrane proteins of lysosomes and endosomes (Marsh *et al*., 1987). Due to their high degree of glycosylation (Fukuda, 1991), lysosomal glycoproteins (LGPs) main function was assumed to be the protection of the lysosomal membrane from the lysosomal hydrolases. However, new roles have arisen, such as the transport of metabolites across the lysosomal membrane (Hogue *et al*., 2002; Kalatzis *et al*., 2001) and chaperone-mediated autophagy (Eskelinen *et al*., 2003).

Since the discovery of the lysosomes, the concept of an acidic membrane-bound organelle, rich in acid hydrolases, that functions as the final destination for substances that enter the cell via endocytosis has evolved to include, not only catabolism, but also other functions such as the turn over of secretory molecules, regulation of surface receptors, inactivation of pathogenic organisms, plasma membrane repair and the loading

of processed antigens onto MHC class II molecules (Mullins and Bonifacino, 2001).

6.1.2 Biogenesis

The process of lysosome biogenesis has been throughly studied and several models have been proposed to describe this mechanism (Luzio *et al*., 2007; Mullins and Bonifacino, 2001). Initial studies suggested the formation of lysosomes from budding directly from the Golgi apparatus. The “maturation” model then proposed that lysosomes would be formed from the maturation of endosomes. In this model, endosomal compartments would form by “pinching” from the plasma membrane and after the addition of *trans*-golgi network (TGN) vesicles be converted to late endosomes and eventually into lysosomes.

Later theories propose that late endosomes and lysosomes are stable, individual compartments, that exchange materials at a certain time point. Either through vesicular transport, multi vesicular bodies, from early endosomes to late endosomes that maturate into lysosomes, or directly to lysosomes (Figure 1).

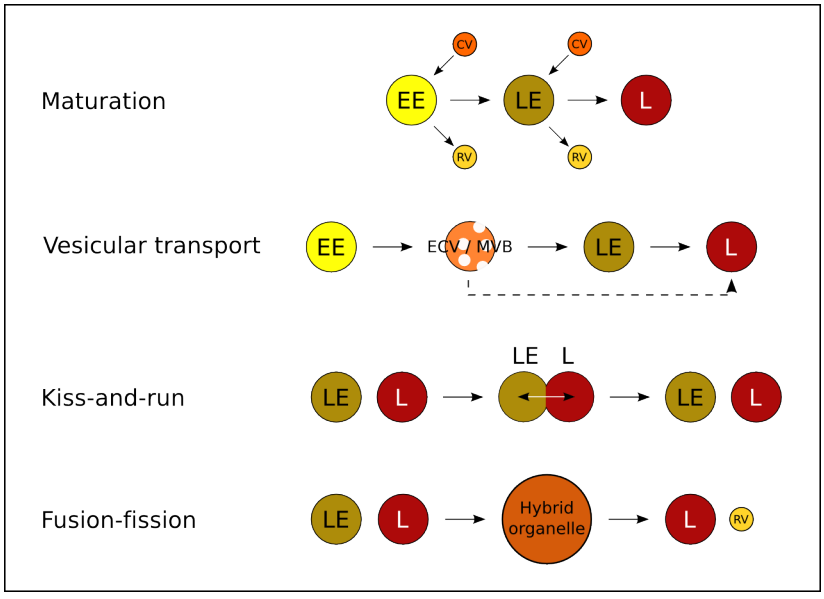


Figure 1. Models for lysosomal biogenesis requiring endosomal compartments. Abbreviations: EE, early endosome; LE, late endosome; L, lysosome; ECV, endosomal carrier vesicle; MVB, multi-vesicular body; CV, cargo vesicle; RV, recycling vesicle. Adapted from Mullins and Bonifacino (2001).

The “kiss-and-run” hypothesis proposes that late endosomes and lysosomes undergo a continuous cycle of transient fusion (“kiss”) and fission (“run”), allowing for the transference of material between both organelles. A modification of this theory, the fusion-fission or “hybrid” model, supports that lysosomes and endosomes fuse permanently and form a hybrid organelle in which material is digested. Lysosomes would then re-form from the hybrid organelle through a process of content condensation and removal of endosomal proteins. This hypothesis was first discussed by Griffiths (Griffiths, 1996), and various data have emerged that document endosomal-lysosomal fusion and the existence of hybrid organelles (Bright *et al*., 1997; Mullock *et al*., 1998; Ward *et al*., 2000). Moreover, recent live-cell imaging experiments have shown that in normal rat kidney cells both fusion and “kissing” events occur, but no vesicle mediated trafficking was observed (Bright *et al*., 2005). Altogether, cell models and cell-free systems have showed that some of the protein machinery involved in these processes has common characteristics to other fusion events in the endocytic and secretory pathways: N-ethylmaleimide sensitive factor (NSF); soluble NSF attachment proteins (SNAPs) and a small GTPase of the RAB family. This, and the mechanistic steps involved in late endosome/lysosome fusion, have been thoroughly reviewed by Luzio (2007).

6.1.3 Trafficking of lysosomal proteins

Correct trafficking of lysosomal proteins produced at the endoplasmic reticulum and Golgi complex is essential to the process of lysosomal biogenesis and to the maintenance of lysosomal function. After arrival at the *trans*-Golgi network, the last sorting location of the biosynthetic pathway, lysosomal proteins can take various pathways to the endosomal/lysosomal system (van Meel and Klumperman, 2008).

6.1.3.1 Mannose-6-phosphate dependent transport of soluble lysosomal proteins

In mammals, the bulk of newly synthesized soluble lysosomal hydrolases is trafficked to the lysosomes via the mannose-6-phosphate receptors (MPRs). After glycosylation in the endoplasmic reticulum, proteins are transported to the Golgi complex where the oligosaccharides are trimmed and the mannose-6-phosphate (M6P) residue is attached. The enzyme N-acetylglucosamine-1-phosphotransferase adds a N-acetylglucosamine-phosphate group to the mannose residues, after which the N-

acetylglucosamine is removed by the N-acetylglucosamine-1-phosphodiester α-N-acetylglucosaminidase (Varki and Kornfeld, 1980; Waheed *et al*., 1981). The moiety is now ready to be recognized by the M6P receptor binding site. Two types of MPRs exist in the majority of mammalian cell types: cation dependent receptor (CD-MPR) - with only one M6P binding site; and the cation independent receptor (CI-MPR) - with two M6P binding sites. Lysosomal proteins can exhibit particular affinities to either CI or CD-MPR (von Figura, 1991). At the TGN, the receptor and the ligand bind and are sorted and packed into clathrin coted vesicles through the action of the adaptor-protein complex 1 (AP1) and GGAs (Golgi localized, γ-ear-containing, Arf-binding family of proteins) (Doray *et al*., 2002). Vesicles traffic to the endo-lysosomal system, directly or via plasma membrane, and after exposure to the relatively low pH of late endosomes, the MPR dissociates from the hydrolase and recycles back to the TGN (Hunziker and Geuze, 1996). The CI-MPR is also involved in the endocytosis of extracellular lysosomal enzymes from the plasma membrane, with the involvement of the adapter-protein complex 2 (AP2) (Rouillé *et al*., 2000).

6.1.3.2 Mannose-6-phosphate independent transport of soluble lysosomal proteins

Several M6P independent trafficking pathways to the lysosomes have been documented. Lysosomal acid phosphatase for example, uses the constitutive secretory pathway to reach the plasma membrane and is then transported to the lysosomes through endocytosis, due to the presence of the tyrosine-based motif YRHV in its cytosolic domain (Peters *et al*., 1990). To date only two receptors independent of M6P have been identified. Lysosomal hydrolase β-glucocerebrosidase, lacking a M6P moiety, is able to bind lysosomal membrane protein LIMP2 in the ER, being routed as a complex to the lysosomes were it dissociates due to the low pH (Reczek *et al*., 2007). It is hypothesized that LIMP2 could be used for the transport of other lysosomal proteins.

Another alternative receptor that has been documented, and plays a role in the trafficking of several lysosomal soluble proteins, including sphingolipid activator proteins (SAPs), is sortilin. Interaction with this 100 kDa multiligand receptor, that cycles between the TGN and the endosomal system, was shown to be necessary for correct lysosomal trafficking of both prosaposin (precursor of saposin A,B,C and D) and G_{M2} activator protein (AP) (Lefrancois *et al*., 2003). The NCL protein cathepsin D was also found to be able to use this alternative sortilin pathway to the lysosomes (Canuel *et al*., 2008b). Another known sortilin ligand is acid sphingomyelinase, involved in a variant form of Niemann-Pick

disease (Ni and Morales, 2006).

Studies in cell types deficient in MPRs or in which the M6P moiety is not added to proteins, such as mouse knockouts or natural occurring disease Mucolipidosis type II (I-cell), respectively, show that at least some lysosomal hydrolases still manage to be correctly targeted to the lysosomes (Dittmer *et al*., 1999; Kornfeld and William, 2001). Thus, these trafficking pathways are not utilized exclusively. Some lysosomal hydrolases that in normal conditions acquire M6P residues can also use these M6P-independent pathways.

6.1.3.3 Transport of lysosomal membrane proteins

The mechanisms of membrane protein transport to the lysosomes is overall still poorly understood. It is known that lysosomal membrane proteins (LMPs) can take either a direct route to the lysosomes after exiting the TGN or in alternative be trafficked via the plasma membrane. Two tyrosine and dileucine motifs (YXXØ, DXXLL), that localize to the cytosolic tails of LMPs, are known to be important for lysosomal sorting (van Meel and Klumperman, 2008). These motifs are present in several lysosomal proteins and have been shown to be involved in both endocytosis at the plasma membrane and direct trafficking from the TGN to the lysosomes. Lamp-1, Lamp-2 and Limp-1, all share the YXXØ consensus motif while Limp-2 for example has the dileucine DXXLL motif (Bonifacino and Traub, 2003). Evidence revealed the existence of both clathrin dependent and independent pathway. Studies showed Lamp-1 sorting into AP1 positive vesicles (Höning *et al*., 1996), but also revealed that depletion of clathrin did not lead to a complete distribution of Lamp-1 at the plasma membrane (Janvier and Bonifacino, 2005).

6.1.4 Lysosomal proteins and brain pathology

The importance of lysosomes in neuronal physiology is evident in the severity of neurodegenerative disorders resulting from defects in lysosomal proteins.

Lysosomal storage diseases (LSDs) are characterized by the accumulation of undegraded enzyme substrates or catabolic products within lysosomes, resulting from the deficient activity of lysosomal proteins. Theoretically, deficiency in any lysosomal protein, not only soluble hydrolases involved in the lysosomal degradative process but also co-factors and transporter proteins for example, could cause an LSD. Over 40 LSDs are

known, a minority being described to involve integral membrane proteins. Most disorders present recessive autossomal inheritance, are normally monogenic and the disease process is usually progressive (Futerman and van Meer, 2004). Classical grouping of LSDs is based on the nature of the accumulated substrates or materials. However, this classical concept has lead to the misclassification of some LSDs, and so a better approach based on the defective enzyme or protein was adopted. Mucolipidosis (ML) II (I Cell disease) and III (pseudo-Hurler polydystrophy) for example, have been classified as lipidosis but are now known to be caused by the defective transport of lysosomal proteins through the mannose-6-phosphate receptor pathway and not by lysosomal lipase deficiency (Kornfeld and William, 2001).

Lysosomal disorders show a diverse clinical phenotype, and may appear as infantile, juvenile or adult forms. Most LSDs present severe neuropathology, with symptoms that may include seizures, dementia and brainstem dysfunction. On the other hand, in some cases such as Fabry disease for example, neuronal involvement is very rare and symptoms are confined to peripheral tissues (Futerman and van Meer, 2004). For each LSD the correspondent clinical and pathological picture should somehow be related to the type of accumulating substrate and the cells and tissues involved. However, most predictions about disease severity and pathology can be inaccurate and even genotype-phenotype correlations can be erroneous, as patients with the same genetic background and mutations can sometimes exhibit very different clinical symptoms. For example in Gaucher disease, which is a sphingolipidosis, neurological involvement should be expected due to the high levels at which glycosphingolipids are present in the brain. Yet, only a small fraction of patients with Gaucher disease show neurological alterations. On the whole, other genes and even environmental factors, as well as affected secondary cellular pathways, may play an important role in defining the disease phenotype and progression.

6.2 Neuronal Ceroid Lipofuscinosis

The Neuronal Ceroid Lipofuscinosis represent a clinical and genetic heterogeneous group of neurodegenerative disorders. The lysosomal etiology was only proposed in the late 90's. The next section summarizes the information gathered through the years and the major contributions to the knowledge on this field.

6.2.1 Common features in NCL

Neuronal ceroid lipofuscinoses are the most common group of inherited neurodegenerative disorders in childhood, with an estimated prevalence between 1:12500 and 1:100000 (Haltia, 2003; Santavuori *et al*., 2000). In Portugal it has been calculated to be 1 per 65000 live births (Teixeira *et al*., 2003). The age of onset of the disease can vary from infancy to adulthood. Early onset forms of the disease manifest through epilepsy, loss of vision and motor and mental deterioration conducting to premature death (Santavuori *et al*., 2000), while the adult form is first associated with intellectual impairment and dementia (Goebel and Braak, 1989).

The term Neuronal ceroid-lipofuscinosis was introduced in 1969 by Zeman and Dyken to distinguish this group of diseases from the gangliosidoses. Its common characteristic feature is the accumulation of an autofluorescence lipopigment, similar to ceroid-lipofuscin, in neuronal cells and other various cell types (Zeman and Dyken, 1969). Based on the nature of the material accumulated in the lysosomes, NCLs can be divided into two groups. While in both late-infantile and juvenile forms subunit c of mitochondrial ATP synthase (SMAS) is the main accumulated proteinaceous component of lipopigments (Palmer *et al*., 1992), infantile and congenital NCL show predominant accumulation of sphingolipid activator proteins (SAPs) A and D (Siintola *et al*., 2006b; Tyynelä *et al*., 1993). The major stored compound seems to be related to the ultra structural morphology of the storage vesicles. Saposins A and D have been associated with GROD (granular osmiophilic deposits), whereas storage of subunit c of ATP synthase has been associated with more variable patterns such as curvilinear, rectilinear and fingerprint profiles (Haltia, 2006).

Before the advent of the molecular genetic era, the age of onset of the disease and the ultra structural morphology of the accumulated deposits were the basis for the classification of NCLs into 4 major classical forms – Infantile (INCL), Late-infantile (LINCL), Juveline (JNCL) and Adult (ANCL) (Zeman, 1976). However, a considerable percentage of NCL cases did not fit into this four classical forms classification (Dyken, 1988). Since 1995, advances in modern molecular biology allowed for the discovery and characterization of several genes underlying NCL. This provided the basis for the current classification, that takes into account the affected gene and permits a better understanding into the molecular pathological mechanisms of this group of disorders.

6.2.2 Genetic heterogeneity

Contrasting with the relatively uniform morphological phenotype described above, a complex genetic heterogeneity underlying NCL has been discovered (NCL mutation database). To date eight different genes associated with the disease have been characterized. Table 1 aims to summarize the current NCL classification, showing identified and postulated NCL genes and associated phenotypes, as well as storage material and gene product.

Table1. The Human neuronal ceroid-lipofuscinosis (NCLs)

Disease	Clinical phenotype	Stored protein	Ultrastruct.	Chromosome	Gene	Protein
CLN1	INCL (LINCL, JNCL, ANCL)	Saposin A, D	GROD	1p32	*CLN1*	PPT1, palmitoyl protein thioesterase 1, soluble
CLN2	LINCL	Subunit c of ATP synthase	CL	11p15	*CLN2*	TPP1, tripeptidyl peptidase 1, soluble
CLN3	JNCL	Subunit c of ATP synthase	FP (CL, RL)	16p12	*CLN3*	CLN3, transmembrane
CLN4	ANCL	Subunit c of ATP synthase	FP, granular	?	?	?
CLN5	$vLINCL_{Finnish}$	Subunit c of ATP synthase	RL, CL, FP	13q22	*CLN5*	CLN5, soluble (?)
CLN6	vLINCL	Subunit c of ATP synthase	RL, CL, FP	15q21-23	*CLN6*	CLN6, transmembrane
CLN7	$vLINCL_{Turkish}$	Subunit c of ATP synthase	RL, CL, FP	4q28	*MFSD8*	MFSD8, transmembrane
CLN8	EPMR	Subunit c of ATP synthase	CL	8p32	*CLN8*	CLN8, transmembrane
CLN9	JNCL	Subunit c of ATP synthase	CL (FP, GROD)	?	?	?
CLN10	Congenital, LINCL	Saposin A, D	GROD	11p15.5	*CTSD*	cathepsin D, soluble
Parry	ANCL (autosomal dominant)	Saposin A, D	GROD	?	?	?

Abreviations: INCL, LINCL, JNCL and ANCL: Infantile, late infantile, juvenile and adult onset neuronal ceroid-lipofuscinosis; EPMR, Epilepsy featuring progressive mental retardation; GROD, Granular osmiophilic deposits; CL, Curvilinear profiles; FP, Fingerprint bodies; RL, Rectilinear profiles.

The first NCL genes to be identified were those associated with the so called classical phenotypes: *CLN1*/INCL, *CLN2*/LINCL and *CLN3*/JNCL. The *CLN1* gene was discovered in 1995 and encodes the soluble lysosomal enzyme palmitoyl protein thioesterase 1 (PPT1) (Vesa *et al*., 1995). In the same year the *CLN3* gene was identified through positional cloning (Lerner *et al*., 1995), coding for a transmembrane protein which has been located to lysosomes (Järvelä *et al*., 1998; 1999), Golgi apparatus (Kremmidiotis *et al*., 1999) and membrane lipid rafts (Rakheja *et al*., 2004). Since then, 6 more genes have been characterized, as being associated with human NCL (*CLN2, CLN5, CLN6, CLN8, CTSD/CLN10 and MSDF8*). NCL genes are found dispersed in the genome and do not occur clustered in a specific chromosomal region.

As is the case with the *CLN1* gene, also *CLN2* (Sleat *et al*., 1997) and *CTSD* (Siintola *et al*., 2006b) code for soluble lysosomal proteins: tripeptidyl peptidase I (TPP1) (Vines and Warburton, 1999); and cathepsin D (Press *et al*., 1960; Rawlings and Barrett, 1995), respectively. On the other hand, both *CLN6* and *CLN8* genes encode transmembrane proteins not localized to the lysosomes, but to the ER and ER-Golgi compartment (Mole *et al*., 2004; Lonka *et al*., 2000). However, while CLN6 is associated with the vLINCL form (Gao *et al*., 2002; Wheeler *et al*., 2002), defects in the *CLN8* gene lead to NCL subtype EPMR (progressive epilepsy with mental retardation) (Ranta *et al*., 1999). The human *CLN5* gene is associated with the Finnish vLINCL and encodes possibly for both soluble and transmembrane lysosomal proteins (Isosomppi *et al*., 2002; Vesa *et al*., 2002), also with unknown function. The most recently identified gene underlying NCL is the *MFSD8/CLN7* gene. It is translated to a novel membrane protein belonging to the major facilitator super family of transporter proteins and was identified in a subset of Turkish vLINCL families (Siintola *et al*., 2007).

Although 8 NCL genes have already been identified so far, at least 3 more putative genes are postulated to exist (*CLN4, CLN9* and "*Parry*"): the *CLN4* gene is proposed to be behind the recessively inherited adult onset form of NCL, the mildest NCL subtype (Berkovic *et al*., 1988); the hypothesized *Parry* gene, responsible for an identified autosomal dominant form of adult NCL (Nijssen *et al*., 2003), and the *CLN9* gene, which leads to a phenotype similar to JNCL, and should encode a regulator of dihydroceramide synthase (Schulz *et al*., 2004; 2006).

An important concept to retain is that different mutations in the same NCL gene can lead to different clinical phenotypes and various onsets of the disease, depending on the location and type of mutation. On the other hand, mutations in different NCL genes can cause very similar clinical phenotypes. Usually, common mutations in a given NCL

gene are usually associated with the “classical” phenotype. The majority of the described mutations correspond to private mutational events and so more mutations are expected to be identified in the future. Mutational analysis and characterization is important, not only at an epidemiological level, but also clinically as it allows for a correct diagnosis of the disease variant, prevention and genetic counseling in the family at risk for a specific NCL variant. Moreover, they are valuable models to analyse functionally important amino acid residues within the protein and aid in the understanding of NCL protein function and biology.

6.2.3 NCL animal models

Animal models are highly valuable to the study of human diseases. They facilitate the study of the underlying pathological mechanisms and represent precious tools for screening and trial of potential therapeutic strategies.

Neuronal ceroid-lipofuscinosis is not exclusive to humans. Even before the identification of the first human NCL gene, several natural occurring diseases, very similar to human NCL, had been reported in other animals (Bildfell *et al*., 1995; Bronson *et al*., 1993; Järplid and Haltia, 1993; Koppang, 1988; Palmer *et al*., 1986). To date, mutations have been described in up to 10 mammalian homologues to human NCL genes (NCL mutation database), that cause NCL in mice, canines, bovines and ovines: mouse *cln6/nclf* and *cln8/mnd* genes (Ranta *et al*., 1999; Wheeler *et al*., 2002); canine homologues of human *CLN8/CLN2/CLN5* and *CTSD* (Awano *et al*., 2006a; Awano *et al*., 2006b; Katz *et al*., 2005; Melville *et al*., 2005); the bovine homologue of human *CLN5* (Houweling *et al*., 2006); and ovine homologues of *CLN5/CLN6* and *CTSD* (Frugier *et al*., 2008; Tammen *et al*., 2006; Tyynelä *et al*., 2000).

Several animal NCL models have also been created through molecular genetic manipulation. Mouse gene knockout models for homologues of most of the known human NCL genes have been developed, and their contribution to the understanding of the biology of NCL has been recently reviewed by Cooper and coauthors (Cooper *et al*., 2006). Other small vertebrates models such as Zebra fish are being developed, while an invertebrate *Drosophila melanogaster* with CLN1 and cathepsin D deficiency (Hickey *et al*., 2006; Myllykangas *et al*., 2005) and a *Caenorhabditis elegans* model for CLN3 (de Voer *et al*., 2005) have also been created. Another well studied model is a *Saccharomyces cerevisiae* strain lacking the human *CLN3* orthologue *Btn1p* (Pearce and Sherman, 1998; Pearce, 2001).

Although these natural occurring and genetic engineered animal models are never a perfect mimic of human NCL, their study has already contributed heavily to the understanding of the pathophysiology of this group of diseases.

6.2.4 Biological function of the NCL proteins

NCL genes encode a variety of proteins with diverse cellular localization and topology. The similarities observed among the NCL disease variants have lead researchers to hypothesize that these proteins may play a role in one common biological pathway particularly relevant to neuronal cells. However, the precise biological function of all NCL proteins is still unknown.

Soluble NCL proteins, such as PPT1/CLN1, TPP1/CLN2 and cathepsinD/CLN10 have been throughly studied. However, their *in vivo* substrates remain elusive. Palmitoyl protein thioesterase 1 (PPT1) removes fatty acids covalently linked to cysteine residues in proteins and *in vitro* studies showed it is able to act upon and remove palmitate from Ras, α subunits of G proteins and palmitoyl-CoA (Camp and Hofmann, 1993; Camp *et al*., 1994). PPT1, although a classical lysosomal enzyme in non-neuronal cells (Hellsten *et al*., 1996; Verkruyse and Hofmann, 1996), has been documented to locate to axons, synaptosomes and synaptic vesicles in neurons (Ahtiainen *et al*., 2003; Heinonen *et al*., 2000a; Lehtovirta *et al*., 2001). This is interesting, as palmitoylation is known to modulate several aspects of neuronal development and synaptic transmission (Huang and El-Husseini, 2005). Cathepsin D (CTSD) is a major lysosomal aspartic protease (Tang and Wong, 1987). *CTSD*-/- knockout mice have revealed that its main activity is related to limited proteolysis of proteins regulating cell growth or tissue homeostasis and not bulk lysosomal proteolysis (Saftig *et al*., 1995). Several proteins have been identified as *in vitro* substrates for CTSD, including prosaposin, the common precursor for saposins A, B, C and D. Saposins are crucial cofactors for the lysosomal degradation of sphingolipids (Kolter and Sandhoff, 2006), and saposins A and D are the main proteolytic accumulated material in both INCL and congenital NCL (Tyynelä *et al*., 1997). TPP1/CLN2 and CLN5 proteins will be reviewed in further detail in a dedicated chapter.

Transmembrane NCL proteins have been more difficult to study and as such are not so well characterized. CLN6 and CLN8 are polytopic transmembrane proteins, that locate to the endoplasmic reticulum (ER) in neuronal and non-neuronal cell types (Heine *et al*., 2007; Lonka *et al*., 2004; Mole *et al*., 2004). Moreover, CLN8 is suggested to recycle between the ER and the ER-Golgi intermediate compartment (Lonka *et al*., 2000).

Although CLN6 is conserved among vertebrates, no sequence or functional homology with other proteins has been found. On the other hand, CLN8 belongs to a large eukaryotic protein family of TLC (Tram-Lag1-CLN8) domain homologues. Members of this family of proteins have postulated functions in lipid biosynthesis, transport and sensing (Winter and Ponting, 2002). Of relevance, roles in the regulation of acyl-CoA dependent ceramide synthesis have also been reported (Guillas *et al*., 2001; Mizutani *et al*., 2005). Interestingly, the so far unknown protein involved in the recently described CLN9 phenotype, is proposed to be a regulator of dihydroceramide synthase and so could be functionally related to CLN8 (Schulz *et al*., 2006). The CLN7 protein is also poorly characterized. It is a postulated transmembrane protein, and was localized to the lysosomal system in transfected COS-1 cells. Homology analysis revealed that CLN7 belongs to the major facilitator superfamily (MFS) of proteins (Siintola *et al*., 2007). MFSs transport specific classes of substrates, including simple sugars, oligosaccharides, inositols, amino acids, nucleosides and a large variety of organic and inorganic anions and cations (Pao *et al*., 1998).

The most comprehensively studied transmembrane NCL protein is definitely CLN3. Its intracellular localization is documented to the lysosomes in non-neuronal cell types and in neuronal cells also to synaptosomes (Järvelä *et al*., 1998; Luiro *et al*., 2001). In addition, it has also been detected at the Golgi complex (Kremmidiotis *et al*., 1999), mitochondria (Katz *et al*., 1997) and plasma membrane lipid rafts (Rakheja *et al*., 2004). Likewise other NCL proteins its precise function is also unknown. However, it has been proposed to participate in multiple cellular processes. Yeast models with a deletion in the *CLN3* homologue *btn1* have proved to be valuable tools. Studies in this model provided evidence for CLN3/btn1 involvement in lysosomal pH homeostasis (Gachet *et al*., 2005) and arginine transport (Kim *et al*., 2003). Palmitoyl-protein Δ-9 desaturase activity has also been reported for CLN3 (Narayan *et al*., 2006). CLN3 overexpression has been shown to enhance growth and suppress apoptosis (Puranam *et al*., 1999) and specific motifs have been identified that are crucial to this behavior (Persaud-Sawin *et al*., 2002). Involvement in endocytic events has also been proposed (Codlin *et al*., 2008; Luiro *et al*., 2001), and recently interactions with plasma membrane associated cytoskeleton and endocytic protein fodrin and associated Na(+), K(+) ATPase have been documented (Uusi-Rauva *et al*., 2008). A recent report has shown that CLN3 protein binds galactosylceramide (GalCer) through a conserved GalCer lipid raft binding domain, proposing therefore a role in the membrane raft transport of GalCer at the ER/Golgi (Rusyn *et al*., 2008).

Little evidence has been reported on direct interactions between NCL proteins. Vesa

and colleagues were the first to describe NCL protein interactions, between CLN5 and CLN2/CLN3, by the use of co-immunoprecipitation and *in vitro* binding assays (Vesa *et al*., 2002). More recently, one report using a similar methodology revealed comprehensive interactions between many NCL proteins, including: CLN3 and CLN6/CLN8, CLN6 and CLN8, CLN2 and CLN1/CLN3/CLN6/CLN8 (Persaud-Sawin *et al*., 2007). Moreover, the same report also proposed functional relations based on the observations that CLN3/CLN6/CLN8 were able to interchangeably correct growth and apoptotic defects in CLN3, CLN8 and CLN6 deficient cells. In addition, CLN2 was able to rescue defects in CLN3, CLN6 and CLN8 deficient cells, albeit only partially for CLN3. This study provides the first experimental data supporting the hypothesis that NCL proteins may interact with each other and play a role along a common biological pathway.

6.2.5 Molecular pathophysiology

It is not yet completely clear which mechanisms are responsible for the selective death of neuronal cells observed in NCL. Apoptosis and autophagy are known mechanisms of cell degeneration that have been reported to be involved in several neurodegenerative conditions, including NCLs (see bellow). It is possible that these processes can occur simultaneously or sequentially, leading to cell death, eventually in a cell-type specific manner. Inflammation and lipid metabolism disturbances are events which have been documented to occur in several NCL subtypes, and are likely to contribute to the disease neurodegenerative phenotype. In the future, a better understanding of these events will be crucial to the development of effective treatments.

6.2.5.1 Apoptosis and autophagy

Apoptosis is a tightly regulated programmed cell death mechanism, known to be involved in the pathophysiology of a variety of neurodegenerative disorders, such as Alzheimer's, Parkinson's and Huntington's disease (Okouchi *et al*., 2007). Likewise, it has been proposed to be one of the principal neuronal cell death mechanisms at least in some NCL subtypes (Mitchison *et al*., 2004). Apoptotic cells have been observed in human LINCL and JNCL CNS tissue (Dhar *et al*., 2002), as well as in PPT1 and cathepsin D mouse models (Gupta *et al*., 2001; Koike *et al*., 2003). Moreover, some NCL proteins have been shown to have anti-apoptotic properties. Of notice, several studies have associated increased PPT1 expression to decreased sensitivity to apoptosis and *vice-*

versa (Kyttälä *et al*., 2006). Apoptosis involvement in JNCL cell death has also been widely documented, and several reports describe anti-apoptotic properties for the CLN3 protein, probably through ceramide level modulation (Persaud-Sawin *et al*., 2002; 2005; Puranam *et al*., 1999).

Autophagy is another key process in cellular homeostasis. Disturbances in autophagic processes have been associated with numerous pathologies, including neurodegenerative diseases and lysosomal storage disorders (Settembre *et al*., 2008; Rajawat and Bossis, 2008). It is also thought to play a role in NCL pathogenesis (Kyttälä *et al*., 2006). Data from a cathepsin D deficient mice, the animal model for congenital NCL, have revealed increased autophagosomes-like bodies accumulation associated with LC3 (light chain 3 of neuronal microtubule-associated protein 1A/B) maturation (Koike *et al*., 2003; 2005), a protein required for autophagosome formation (Kabeya *et al*., 2000). Autophagy involvement could provide a possible explanation for the accumulation of subunit c of mitochondrial ATP synthase in the majority of NCL subtypes.

6.2.5.2 Inflammation

Lysosomes play a very important role in several functions required for normal immune system function. Among others, it is responsible for antigen presentation via major histocompatibility (MHC) class II molecules, cytotoxic T cell function and lipid presentation by CD1d molecules in the context of NKT cell development. Recent studies have revealed links between neuroimmune responses and CNS pathology in some lysosomal storage disorders, including neuronal ceroid lipofuscinosis (Castaneda *et al*., 2008; Cooper, 2003). Transcript expression profiling performed on tissues from PPT1 and CLN5 mouse models have revealed an increased expression in inflammatory associated genes, suggesting an augmented inflammatory response (Jalanko *et al*., 2005; Kopra *et al*., 2004). cDNA microarray studies on vLINCL/CLN6 human fibroblasts samples suggest changes in both pro and anti-inflammatory molecules (Teixeira *et al*., 2006). Moreover, there is mounting evidence for the activation of glial cells in NCL, as astrocytosis and microgial activation have been thoroughly documented in some NCL subtypes, including CLN3, CLN6 and congenital NCL (Castaneda *et al*., 2008; Jalanko and Braulke, 2008). In addition to these immune mechanisms, wide range autoimmune response has also been observed for CLN3 patients, shown by elevated IgG levels in the CNS (Lim *et al*., 2006). However, it is still unclear how and to what degree these processes contribute to the pathology observed in NCL patients.

6.2.5.3 The role of lipids

Evidence suggesting disturbances in lipid metabolism have been presented for several NCL subtypes. Moreover, some NCL proteins are speculated to have direct interaction with lipids or act in lipid related processes (Jalanko and Braulke, 2008).

Transcript profiling analysis on *PPT1* knock-out mice has shown deregulation of important neuronal functions, including cholesterol metabolism. Genes related to cholesterol biosynthesis, including components of the HMG-CoA reductase pathway were found to be strongly up-regulated, along with the observation of reduced amount of steady state cholesterol precursors (Ahtiainen *et al*., 2007). Another report has identified a direct interaction of PPT1 with the F(1)-complex of the mitochondrial ATP synthase. F(1)-complex subunits are documented to localize also at the plasma membrane and act as receptors for apolipoprotein A-1. In fact, strong changes were detected in the apolipoprotein A-1 uptake of *PPT1* mouse knockout neurons and in serum lipid composition (Lyly *et al*., 2008). Additionally, loss of phospholipid content and alterations in the phospholipid composition of membranes has been observed in the cerebral cortex of LINCL patients (Käkelä *et al*., 2003).

Cholesterol related alterations were also found in vLINCL/CLN6 patients. Gene expression analysis revealed altered profiles for genes connected to cholesterol as precursor for steroid hormones, structural component of lipid rafts and cofactor for signaling. Moreover, free cholesterol and ganglioside GM2 were found to accumulate in the lysosomal compartment (Teixeira *et al*., 2006). Similar findings have been recently described for cathepsin D deficient mice. Increased levels of GM2 and GM3 gangliosides were observed, along with lysophospholipid bis(monoacylglycero)phosphate and mild unesterified cholesterol accumulation (Jabs *et al*., 2008). Cathepsin D is proposed to play an important role in sphingolipid and cholesterol intracellular trafficking. It is also necessary in the proteolytic processing of prosaposin, precursor of saposins (A-D), which are required for the hydrolysis of sphingolipids and important for the degradation of ceramide. Moreover, CTSD has been shown to be involved in the regulation of ABCA1-mediated cholesterol eflux (Haidar *et al*., 2006).

A recent and comprehensive report has characterized altered sphingolipid composition of membrane lipid rafts on CLN3 deficient cells. Furthermore, the authors showed the CLN3 protein binds galactosylceramide (GalCer) through a conserved GalCer lipid raft binding domain, and so have proposed a role in the ER/Golgi to membrane raft transport of GalCer (Rusyn *et al*., 2008). This analysis was extended to other NCL

proteins and revealed alterations in raft sphingolipid composition for CLN1/CLN2/CLN6/CLN8/CLN9 deficient cells, along with interaction of CLN6/CLN8 with GalCer, CLN8 with ceramide and CLN1/CLN2/CLN6 with lysophosphatidic acid/sulfatide.

The CLN8 and CLN9 proteins also seem to have lipid related functions. Disturbances in sphingolipids metabolism have been detected in brain samples from CLN8 patients as well as in CLN9 deficient cells, and include reduced levels of several sphingolipids species such as ceramide, galactosyl- and lactosylceramide, sphingomyelin and globoside (Hermansson *et al*., 2005; Schulz *et al*., 2004). CLN8 shares functional homology to the TLC family of proteins, which includes yeast Lag1 and human homologue LASS1, known to regulate dihydroceramide synthesis (Riebeling *et al*., 2003). A similar function is hypothesized for the currently unidentified CLN9 protein, since it has been shown that Lag1/LASS1 are able to rescue ceramide levels in CLN9 deficient cells (Schulz *et al*., 2006).

Overall, published data suggest that cholesterol and sphingolipid homeostasis disturbances are a common theme for NCL. De-regulation of these processes can lead to alterations in membrane composition and to the disruption of signal transduction cascades and vesicular trafficking, contributing to the disease progression and severity.

6.2.6 Therapeutic Strategies

At the moment, there is no available treatment for neuronal ceroid-lipofuscinosis. Genetic characterization of NCLs has started only in the last 15 years and the pathological mechanisms causal to disease are still poorly understood for most NCL variants. However, for some better characterized forms of NCL, associated with the deficiency in soluble proteins, particularly infantile (CLN1) and late-infantile (CLN2), several strategies are undergoing testing and evaluation. Approaches that have been utilized in other lysosomal storage diseases with relative success, such as enzyme replacement therapy (ERT), gene and cell mediated therapy, and treatment with pharmacological molecules, have been applied to several animal models of NCL and some are even going through the first stages of clinical trials (Pierret *et al*., 2008).

ERT for TPP1/CLN2 has very recently been applied to a mouse model of LINCL, through intra ventricular delivery of the enzyme, with reports of attenuated neuropathology, hence demonstrating potential therapeutic value (Chang *et al*., 2008). Positive results were also obtained with gene therapy through adeno-associated virus 2/5 (AAV2/5) delivery of TPP1 to CLN2 mouse models and non-human primates (Griffey *et*

al., 2004; Passini *et al*., 2006; Sondhi *et al*., 2005). This has prompted application to human LINCL patients (Worgall *et al*., 2008) and a Phase I clinical trial is currently underway to access safety of the gene vectors (Crystal, 2007). A similar strategy has been applied to PPT1/CLN1 and a mouse knockout model for INCL, which resulted in reduced lysosomal storage and milder phenotype (Griffey *et al*., 2004). Cell-mediated therapy in the form of stem cell transplant is also very promising. Two clinical trials are currently taking place to evaluate the safety and engraftment of donor hematopoietic cells in CLN3 patients (Orchard, 2007), and to study the safety and effectiveness of human CNS stem cell transplant in CLN1 and CLN2 patients (StemCells, 2006). The use of pharmacological agents is another method being tried on NCLs. A combined treatment with Cystagon (cysteamine) and Mucomyst (N-acetylcysteine), shown to reduce substrate deposits in INCL patients (Hobert and Dawson, 2006), has entered Phase II clinical trial with INCL patients (NICHD, 2007).

In this regard, knowledge on the function of the proteins encoded by NCL genes and a deep understanding of the pathological metabolic mechanisms behind neuronal ceroid-lipofuscinosis will undoubtedly contribute to the effectiveness of current and future therapeutic approaches.

6.3 Late-infantile NCL: CLN2 and CLN5

6.3.1 Clinical phenotype and neuropathology

6.3.1.1 CLN2

Patients with the late-infantile onset of NCL (OMIM: 204500) were first documented in 1908 by Janský and later by Bielschowsky in 1923, hence LINCL being also referred to as Janský-Bielschowsky disease (Bielschowsky, 1913; Janský, 1908). LINCL, as classically defined, results from mutations in the *CLN2* gene. In the typical onset, symptoms appear at the age of 2-4 years, usually starting with clumsiness and epilepsy (tonic-clonic and/or secondary generalized seizures), followed by progressive loss of motor and cognitive skills. Visual failure leading to complete loss of sight also develops, with death following in few years. However, some mutations in the *CLN2* gene have been

associated with a slower progression of symptoms as they were observed in patients with juvenile and even later onsets of the disease (Sleat *et al*., 1999; Steinfeld *et al*., 2002; Zhong *et al*., 2000).

Classical late infantile NCL features include massive neuronal loss in the cerebral and cerebellar cortices. Neuron damage can vary, with the co-existence in the same area of the brain of both structurally intact cells and heavily damaged neuron showing huge amount of lysosomal storage (Wisniewski *et al*., 1992). Another observed anomaly is the diffuse proliferation of macrophages and astrogliosis (Brück and Goebel, 1998). Deposition of storage materials is observed in neurons but also in many tissues and organs. The main component of the lysosomal storage is the subunit c of the mitochondrial ATP synthase (Palmer *et al*., 1989; 1992), although some neurons in certain areas of the brain show inclusions of unclear origin, that are negative for both SCMAS and SAPs (Kida *et al*., 1995). Curvilinear profiles are the predominant ultrastructure appearance of storage deposits, even though in some cases mixed patterns are observed (Wisniewski *et al*., 1988).

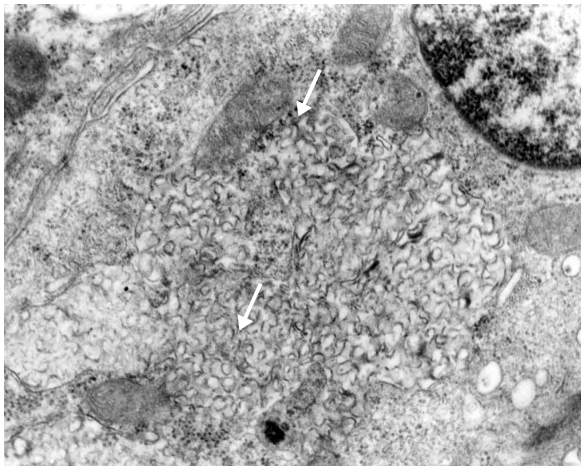


Figure 2. Ultrastructure of a muscle biopsy from a LINCL patient. Curvilinear inclusions are noted with white arrows. Adapted from Teixeira *et al*. (2003).

6.3.1.2 CLN5

The Finnish variant late infantile NCL ($vLINCL_{Fin}$) (OMIM: 256731) was first described in 1991 (Santavuori *et al*., 1991), and results from mutations in the *CLN5* gene

(Savukoski *et al*., 1998). Up to very recently it was confined to Finland. These cases are very similar to the classic late-infantile NCL with regard to clinical phenotype, presenting clumsiness and muscular hypotonia at first, advancing into developmental regression, blindness, epilepsy and eventually death. A later onset of the disease and different EM profiles differentiates Finnish late infantile variant subjects from the classic late infantile phenotype. A recent report has described a novel *CLN5* mutation in a Colombian family associated with an atypical milder phenotype (Pineda-Trujillo *et al*., 2005).

All $vLINCL_{Fin}$ patients show severe generalized cerebral atrophy, particularly acute in the cerebellum, and visible ventricular dilatation. Cell death is not observed outside the CNS. Histologically, most nerve cells present storage granules in the cytoplasm. The same storage can be observed in other cell types, but to a lesser amount. The major compound of the storage material is the subunit c of the mitochondrial ATP synthase, but a minor accumulation of saposins A and D is also detected. Ultrastructure of storage deposits comprises rectilinear complexes and both fingerprint and curvilinear patterns (Santavuori *et al*., 1982; Tyynelä *et al*., 1997b).

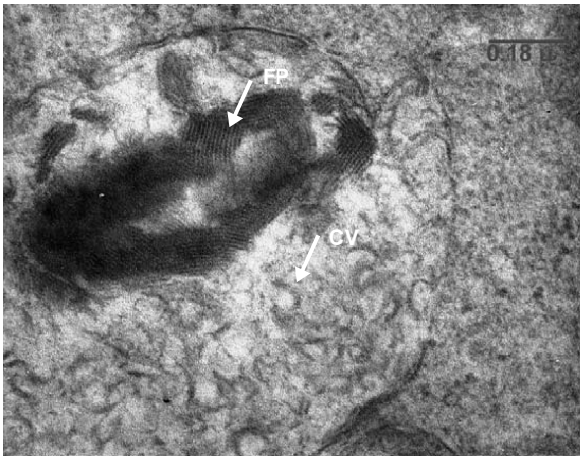


Figure 3. Ultrastructure of a muscle biopsy from a variant LINCL patient. Curvilinear and fingerprint profiles are noted with arrows (CV/FP). Adapted from Bessa *et al*. (2006).

6.3.2 Genes and mutations

6.3.2.1 Classical Late-Infantile NCL

The *CLN2* gene (MIM: 607998) consists of 13 exons, spanning a coding region of 2612 bp, and is located on the chromosomal region 11p15 (Liu *et al*., 1998; Sharp *et al*., 1997). Two transcripts have been detected by Northern Blot analysis, resulting from the use of alternative polyadenylation signals (Sleat *et al*., 1997). *CLN2* expression is ubiquitous in the CNS and extra cerebral organs, showing correlation with cerebral cortex development (Kurachi *et al*., 2001).

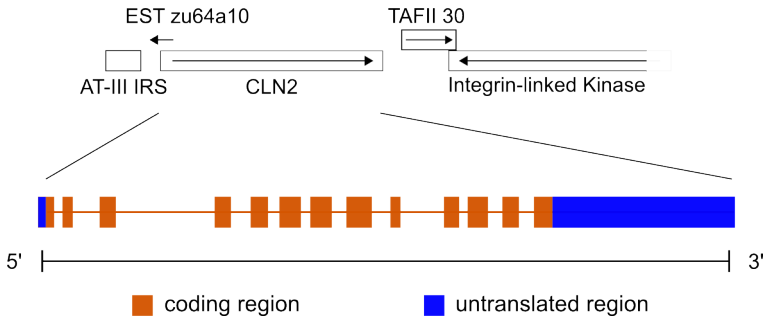


Figure 4. *CLN2* gene structure. Represented here is the position of *CLN2* relative to other proximal genes and its intron:exon structure. Adapted from Liu *et al*. (1998), and the NCBI database (MIM:607998).

To date, 68 mutations and 22 polymorphisms have been described in this gene (NCL mutation database). There are two most common mutations worldwide. The nonsense mutation p.R208X (c.622C>T) (Sleat *et al*., 1997), from which no resulting protein has been detected (Steinfeld *et al*., 2004) and an intron change mutation, IVS5 -1G>C (g.3556G>C) (Sleat *et al*. 1997), which results in the insertion of intron 5, causing a frameshift and leading to the production of a truncated non-functional protein (Hartikainen *et al*., 1999). Most of the other mutations correspond to private mutational events that lead to PPT1 activity deficiency. Mutations that lead to premature stop codons are uniformly associated with severe clinical progression and result in premature protein degradation, misfolding or disrupted lysosomal targeting. In the case of missense mutations a distinct

impact on the protein stability, targeting or impairment of functionally/structurally important protein domains may explain the clinical heterogeneity. Although the majority of the mutations in the *CLN2* gene cause classical late infantile NCL, a few described mutations can result in a later, juvenile, onset of the disease (Siintola *et al*., 2006a).

6.3.2.2 Finnish Late-Infantile Variant NCL

The *CLN5* gene (MIM: 608102) was localized to chromosomal region 13q21.1-q32 and consists of four exons, spanning a coding region of 1221 bp (Savukoski *et al*., 1998). The human *CLN5* gene shows ubiquitous expression throughout several tissues, and four transcripts (2.0, 3.0, 4.5 and 5.5 kb) were detected by Northern hybridization (Savukoski *et al*., 1998). This suggests that *CLN5* may be subjected to tissue or cell-type specific alternative splicing or polyadenylation. *CLN5* expression was found to occur at the beginning of cortical neurogenesis, increasing along with cortical development, and so may be relevant to neuron development and maturation (Heinonen *et al*., 2000b). Similar observations have been reported for the mouse homologue (Holmberg *et al*., 2004).

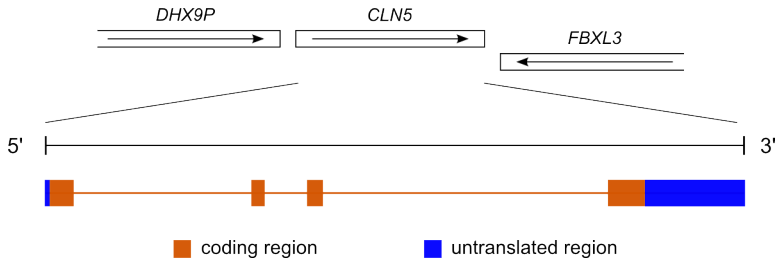


Figure 5. *CLN5* gene structure. Represented here is the position of *CLN5* relative to other proximal genes and its intron:exon strucuture. Adapted from the NCBI database (MIM:608102).

So far, 18 mutations and 1 polymorphism have been found in patients from different countries (NCL mutation database). The majority of cases has been identified in the northern European population but in recent years several new mutations have been identified worldwide. The most frequent mutation, $CLN5_{Fin\ major}$, is a 2 base pair deletion in exon four (c.1175delAT) and results in the substitution of tyrosine to a STOP codon at position 392 (p.Y392X), leading to a predicted truncated protein (Holmberg *et al*., 2000;

Savukoski *et al*., 1998). The $CLN5_{Fin\ minor}$ mutation, the second most common gene defect among $vLINCL_{Fin}$ families, is a base pair substitution (c.225G>A) which creates a STOP codon at position 75 (p.W75X), resulting in a predicted truncated polypeptide of only 74 aminoacids (Savukoski *et al*., 1998). Recently, two new mutations were found in atypical vLINCL patients from Argentina and Colombia. The missense mutation p.R112H (c1627G>A), was associated with a later, juvenile onset of NCL (Pineda-Trujillo *et al*., 2005), while an 1 bp insertion in exon 1 (c.291insC; p.Ser98fs) led to an early onset of the disease (Cismondi *et al*., 2008). Genotype/phenotype correlations are difficult to establish since the function of the protein is unknown and no extensive characterization of the molecular consequences of mutations has been performed. Previous to the work described in this thesis, no mutations had been described in the Portuguese population.

6.3.3 The proteins

6.3.3.1 TPP1

Tripeptidyl peptidase 1 (TPP1) is the defective protein underlying Late Infantile NCL and was identified by a proteomic approach as the missing enzyme in LINCL patient brain samples (Sleat *et al*., 1997). TPP1 is a lysosomal hydrolase that removes peptides from the N-terminus of small polypeptides with an optimum pH of ≈ 5.0 (Golabek and Kida, 2006) and belongs to the family of pepstatin A-insensitive serine carboxyl proteases, also known as sedolisins (Wlodawer *et al*., 2003). Additionally, a much weaker endoproteolytic activity with an optimum pH of ≈ 3.0 is also present (Ezaki *et al*., 2000b). It is synthesized as a 66 kDa precursor, that suffers signal peptide co-translational removal in the ER. Upon arrival at the lysosomal lumen the precursor is converted to a 46 kDa mature enzyme by a process of autocatalytic cleavage due to the low pH (Golabek *et al*., 2003; 2004; Lin *et al*., 2001). Recently, evidence has emerged that suggests inhibitory properties for the 176 amino acid cleaved prosegment torwards its parent enzyme (Golabek *et al*., 2008).

Human TPP1 possesses five potential N-glysolysation sites at Asn residues 210, 222, 286, 313, and 443, that have been shown to be used to their full extent with high-mannose and complex type oligosaccharides (Golabek *et al*., 2003). These modifications were found to be required for both stability and correct folding, but also for proper

lysosomal targeting (Wujek *et al*., 2004) which occurs via mannose-6-phosphate receptor mediated pathway (Golabek and Kida, 2006). Several studies employing homology analysis, mutational and inhibition assays have characterized TPP1 catalytic mechanism. The amino acid Ser475 was identified as the active site nucleophile, that together with Asp276 and Glu272 form the TPP1 catalytic triad (Lin *et al*., 2001; Oyama *et al*., 2005; Walus *et al*., 2005). Very recently, two studies have determined the crystallographic structure of the TPP1 precursor form (Guhaniyogi *et al*., 2009; Pal *et al*., 2009). Combinational peptide libraries have been used to characterize TPP1 substrate specificity, but although some broad rules regarding preferred amino acids could be established no consensus sequence could be determined (Tian *et al*., 2006).

In vivo substrates for TPP1 are unknown, but *in vitro* studies have documented that it can act upon several molecules. It has been shown to be involved and required for the degradation of subunit c of mitochondrial ATP synthase, the major protein component of the storage material in LINCL and several other NCL forms (Ezaki *et al*., 1999; Kominami, 2002). TTP1 is proposed to initiate SCMAS degradation in the lysosome (Ezaki *et al*., 2000a). Additionally TTP1 was described as playing a predominant role in neuromedin B (Du *et al*., 2001; Kopan *et al*., 2004) and cholecystokinin degradation (Bernardini and Warburton, 2002; Warburton and Bernardini, 2002). ß-amyloid and angiotensins II/III have also been proposed as candidates for possible natural TPP1 substrates (Du *et al*., 2001; Junaid *et al*., 2000). In this regard, as NCL is characterized by a pronounced CNS involvement, neuropeptides are attractive candidates for TPP1 substrates.

6.3.3.2 CLN5p

The human *CLN5* gene encodes a 407 amino acid polypeptide with a predicted molecular weight of 46 kDa. CLN5 shares no homology to previously reported proteins and is moderately conserved in mammals (Savukoski *et al*., 1998). It possesses a unique N-terminal extremity with the presence of four initiation methionines (Figure 6) that have been shown to be used both *in vitro* and *in vivo*, resulting in the production of four 39-47 kDa polypeptides (Isosomppi *et al*., 2002; Vesa *et al*., 2002). Initial "*in silico*" predictions, using the most upstream 5' initiation codon suggested the production of a polypeptide with two putative transmembrane domains. However, sequence alignments using mammalian CLN5 homologues seem to point to either the third or fourth ATG initiation methionine as being evolutionary conserved (Holmberg *et al*., 2004). CLN5 contains eight potential N-glycosylation sites. In over expression studies it has been detected as a highly

glycosylated 60 kDa protein with both high mannose and complex type sugars (Isosomppi *et al*., 2002; Vesa *et al*., 2002).

Figure 6. Schematic presentation of the CLN5 polypeptide. Potential initiation methionines are indicated in the upper part, by arrows. Positions of disease mutations are indicated by arrows in the lower part, and putative N-glycosylation sites by stars. Green stars represent experimentally documented N-glycosylation sites.

The solubility of the protein remains controversial. Initial predictions proposed an integral membrane topology, but evidence for both membrane association and soluble species have been documented (Isosomppi *et al*., 2002; Vesa *et al*., 2002). Moreover, overexpressed mouse Cln5 has been shown to be a soluble lysosomal protein and to also localize to axons in neurons (Holmberg *et al*., 2004). Lysosomal localization is supported by studies with immunofluorescence microscopy and, more recently, by proteomics analysis. The CLN5 protein has been isolated from human brain, plasma and urine samples using M6P affinity assays (Sleat *et al*., 2005; 2006a; 2007). A similar approach has confirmed M6P phosphorylation sites in Asn320, Asn330 and Asn401 of the CLN5 protein purified from human brain samples (Sleat *et al*., 2006b). Additionally, CLN5 was found to be secreted by MPR deficient embryonic mouse fibroblasts (Kollman *et al.,* 2005). Overall, these data argues in favor of M6P receptor mediated trafficking of CLN5 to the lysosomes. Co-immunoprecipitation and *in vitro* binding assays have revealed an interaction of CLN5 with two other NCL proteins, CLN2 and CLN3. Data suggests interaction occurs with the longest, membrane associated, CLN5 polypeptide (Vesa *et al*., 2002). The function of the CLN5 protein is currently unknown. However, gene expression profiling analysis using Cln5 knockout mice brain samples has provided some insightful clues. Significant expression alterations in genes related to inflammation and myelination processes were observed (Kopra *et al*., 2004b; Schantz *et al*., 2008). The most significant CLN5 specific finding however, was differential expression of genes related to RNA processing and transcription (Schantz *et al*., 2008). Surprisingly this later study also revealed a “common pathway” with CLN1/PPT1, as common changes in groups of genes implied in phosphorylation signaling cascades, cytoskeleton and growth cones in neurons

were observed in both Cln5-/- and Cln1-/-. So far, these studies provide the only clues to unravel the function of the CLN5 protein.

7 Aims of the Study

Neuronal ceroid lipofuscinoses (NCLs) represent a group of children's recessively inherited neurodegenerative disorders caused by mutations in at least eight different genes. Despite the genetic heterogeneity the clinical findings are highly similar among the genetic variants. Therefore, the identification of the disease variant and the comprehension of the impact of mutations on the phenotype are not always straightforward. As a matter of fact, at the time this project was initiated few NCL proteins had been extensively studied and for some variants even the genetic analysis was scarce. Furthermore, the knowledge on the molecular mechanisms behind gene defects was limited and in most cases restricted to prevalent mutations observed in some disease variants.

With the goal to provide a better understanding on the molecular pathophysiology of this group of genetically heterogeneous neurological diseases this study aimed to:

- Investigate the primary molecular defects of the late infantile form of the disease associated with the deficiency of CLN2 and CLN5 proteins.
- Extend the knowledge on the molecular and cellular properties of CLN5.
- Evaluate the impact of *CLN5* mutations on the cell processing and intracellular trafficking of the protein, and on the functional connection between CLN5 and other CLN proteins, using a novel CLN5-specific antibody.

8 Materials and methods

This chapter describes the materials and methods used in the studies presented in this thesis.

8.1 Patients and clinical phenotype

LINCL patient

The case here reported had onset of symptoms at 10 years of age and presented a progressive cognitive and motor dysfunction, epilepsy and seizures. At the time the study was published the patient was 40 years old.

vLINCL$_{Fin}$ patient

The case here reported had onset of symptoms at 3 years of age. By that time progressive attention deficits and regression of language capacities were noticed with motor clumsiness appearing one year later. At 7 years of age she initiated visual deterioration, gait ataxia, myoclonus (distal and axial), and generalized epileptic seizures (atonic and myoclonic). The neuro-ophthalmologic examination disclosed bilateral optic atrophy and retinal degeneration. The EEG showed diffuse slowing of the background rhythm and occipital spike-wave activity. Brain MRI showed cerebral and cerebellar atrophy. At 8-years-old she became completely dependent in the daily activities, having lost all language abilities and presented a marked ataxic gait. The EEG activity became slower and disorganized with frequent multifocal high-voltage polyspike-slow wave complexes; the video-EEG disclosed that most of the epileptic activity corresponded to myoclonus. At the time the study was published she was 15 years old, bedridden and presented an amaurotic behaviour, emotional instability, and marked swallowing difficulties. The neurological examination disclosed a severe mental retardation, limb and trunk dystonic postures, pyramidal signs in the four limbs, bilateral optic atrophy, and frequent erratic and stimuli-induced myoclonus.

8.2 Biological samples and cell culture

Leukocytes were isolated from peripheral blood collected in vaccutainer tubes containing EDTA-K3 according to a standard procedure described by Skoog *et al*. (1956)

and stored at -70ºC. Human skin fibroblasts, COS-1 and HeLa cells were cultured in Dulbecco´s modified Eagle´s medium (DMEM, Gibco), supplemented with 10 % of fetal calf serum and antibiotics. Cell lines originating from NCL patients were available at the laboratory and their use in NCL research was approved by the ethic committee. Human neuroblastoma cells (SH-SY5Y) were cultured in 1:1 DMEM and Ham´s F12 medium (Sigma), supplemented with 10 % fetal calf serum, antibiotics and 0.1 % non-essential amino acids. Cultures were maintained at 37 °C, in an atmosphere containing 5 % CO_2.

8.3 Electron microscopy

A muscle fragment was fixed in a 4% glutaraldehyde solution in cacodylate buffer (pH 7.4), subsequently fixed in osmium tetroxide and embedded in Epon. Sections of 1 µm were stained with toluidine blue. Fine sections were observed under transmission electron microscope after staining by uranyl acetate and lead citrate. Dr. António Guimarães is acknowledged for the Electron microscopy analysis and imaging.

8.4 In vitro enzymatic assays

Cell pellets were lysed in water by sonication. Lysates were centrifuged for 30 min at 15,000*g* and the supernatant, or extract, was assayed for total protein content through the Lowry method (Lowry *et al*., 1951). The enzymatic activity of palmitoyl-protein thio-esterase 1 (PPT1) and tripeptidyl peptidase 1 (TPP1) was determined using synthetic fluorogenic substrates, respectively, 4-methylumbelliferyl-6-thiopalmitoyl-β-D-glucopyranoside (4MU-6S-Palm-βGlc, Moscerdam Substrates, The Netherlands) and Ala-Ala-Phe-7-amido-4-methylcoumarin (Ala-Ala-Phe-AMC, Sigma, USA), as previously described (van Diggelen *et al*., 1999; Vines and Warburton, 1999). The fluorescence was measured on a Hitachi F2000 spectrofluorometer. Standard solutions of 4MU-6S-Palm-βGlc and Ala-Ala-Phe-AMC were used to calculate the specific enzymatic activity.

8.5 DNA sequencing

Genomic DNA used for the screening of mutations in the *CLN2* gene was extracted from peripheral blood leucocytes with the “Puregene DNA purification kit’” (Gentra systems, USA), while the genomic DNA used for the screening of mutations in the *CLN5*

gene was extracted from the buffy coat of nucleated cells according to standard procedures (Miller *et al*., 1988). All thirteen exons of the *CLN2* gene and four exons of the *CLN5* gene were amplified using intron-based primers (gene sequence Accession No. AF039704 and NG009064, respectively). The amplification reaction was performed with Taq DNA Polymerase from MBI Fermentas (Canada) in UnoII thermocyclers (Biometra, Germany). Reactions started with a denaturing step of 5 minutes at 95 °C, followed by 35 cycles of denaturing (45 seconds, 95 °C), annealing (45 seconds) and extension (45 seconds, 72 °C), and a final extension step of 10 minutes at 72 °C. PCR products were sequenced in a 310 Genetic analyser automatic sequencer (Applied Biosystems, USA), according to the manufacturer's instructions. Mutation nomenclature is in accordance with the guidelines of Human Genome Variation Society (http://1www.hgvs.org/mutnomen).

DNA sequencing was also performed for the *CLN2* cDNA. Total RNA was isolated from human skin fibroblasts using the ‘‘RNeasy Midi kit’’ (Qiagen) or ‘‘Roche Applied Science Kit’’ (Roche), according to the manufacturer’s instructions. The reverse transcription of RNA was performed using the ‘‘First-strand cDNA synthesis kit’’ (Amersham Biosciences) following the manufacturer’s protocol for rare mRNA’s. PCR amplified cDNA was analysed by direct sequencing of four overlapping fragments using the following forward and reverse primers: CLN2F1-5'- TGA CAG CAG ATC CGC GGA A-3' and CLN2R1-5'-GCC TCA GGG ATG ATG TTG G-3' for exons1–4; CLN2F2-5'- TAA GGT CCC CAC ATC CCT ACC-3' and CLN2R2-5'-GCT GAG GGA GTC CTC ATC ATC-3 ' for exons4–7; CLN3F3-5'-GCT GCT CAG TAA TGA GTC AG-3' and CLN2R3-5'-ACC CAG TAG CCA TCA GAA AG-3' for exons7–10; CLN2F4-5'-CAC CTG CCA CCA TCC AGT T-3' and CLN2R4-5'-TGA GGG TTC AGC AGG GCT T-3' for exons10–13. PCR products were sequenced in a 310 Genetic analyser automatic sequencer (Applied Biosystems, USA), according to the manufacturer's instructions.

8.6 Mutation screening

The c.335G>C (p.R112P) mutation was screened by restriction enzyme analysis through the digestion of amplified PCR fragments of *CLN5* exon2 with the enzyme MnI I (New England Biolabs, USA) according to the manufacturer's instructions. The restriction fragments were separated by electrophoresis on a 3 % Nusieve 3:1 agarose (Cambrex, USA) run at 75 V for 30 min. The DNA fragments were visualized by staining with ethidium bromide.

The c.835G>A (p.D279N) mutation was screened by single strand conformational

polymorphism (SSCP) analysis. Amplified PCR fragments of *CLN5* exon4 were heat denatured and loaded onto a 1x MDE gel (Cambrex, USA) prepared in 0.6x Tris/Borate/EDTA buffer (TBE). Electrophoresis occurred at 550 V and 4 ºC for 20 hours. Conformational polymorphisms were visualized by silver staining.

The IVS7-10A>G mutation was screened by heteroduplex analysis. Total RNA was isolated from $5x10^6$ to $1x10^7$ cells using the ‘‘High Pure RNA Isolation kit’’ (Roche, Switzerland) according to the manufacturers instructions. A segment encompassing the exons 4–7 of the *CLN5* mRNA were amplified through RT-PCR with the “First-strand cDNA synthesis kit” (Amersham Biosciences, UK), following the manufacturer's protocol. RT-PCR samples were heat denatured and kept at room temperature for 30 min to allow for the reannealing of the strands (heteroduplexes formation). Samples were loaded onto a 5 % polyacrylamide gel (29:1) prepared in 1x TBE, and separated by electrophoresis at 175 V for 7 hours at room temperature. The pattern of the RT-PCR products was visualized by ethidium bromide staining. Alternatively, RT-PCR products were analysed by electrophoresis at 80 V for 6 h at room temperature, in 3 % Nusieve (Cambrex, USA) agarose gel.

8.7 Northern blot analysis

RNA was isolated from skin fibroblast cell cultures of patients and controls using the ‘‘RNeasy Midi kit’’ (Qiagen, Germany) or ‘‘High Pure RNA Isolation kit’’ (Roche, Switzerland) according to manufacturer’s instructions. Integrity of total RNA was checked by electrophoresis, through the estimation of the 28S/18S ratio. Approximately 25 µg of total RNA was separated by electrophoresis in a 1 % agarose/formaldehyde denaturing gel and afterwards transferred to a nylon membrane (NEN Research products, USA). The oligonucleotides used to generate the cDNA probes were as follows: *CLN2* forward: 5'-TCCCTGGGGCTGAGTTTCA-3'; *CLN2* reverse: 5'-AAAGGTTCCTGGAAGGATGTG-3'; *GAPDH* forward: 5'-CCATCAATGACCCCTTCATTGA-3'; and *GAPDH* reverse: 5'-GAAGGCCATGCCAGTGAGCTT-3'. Hybridization with the normal *CLN2* (nt 427–1219) cDNA probe, labeled with [α-^{32}P]dCTP (RadPrime DNA Labeling System, Invitrogen, USA), was performed with the ‘‘ExpressHyb” hybridization solution (Clontech, Japan), according to the manufacturer's instructions. Blots were exposed at -70 °C and revealed. The *GAPDH* cDNA was used as control probe.

8.8 Relative quantitation of mRNA by real-time PCR

Total RNA was isolated from $5x10^6$ to $1x10^7$ cells using the ‘‘High Pure RNA Isolation kit’’ (Roche, Switzerland) according to the manufacturers instructions. The reverse transcription reaction was performed with the “First-strand cDNA synthesis kit” (Amersham Biosciences, UK) following the manufacturer's protocol. Real-time PCR analysis were performed on an ABI PRISM 7000 Sequence Detection System from Applied Biosystems (USA). PCRs were prepared using 1x TaqMan Universal PCR Master Mix (Applied Biosystems, USA), 300 nM of each primer, 250 nM of probe and 1 ng of RNA converted to cDNA (0,5 ng for the *CLN2*/mRNA in Publication-I), in a final volume of 50 µL. Thermal cycling conditions included an initial AmpErase UNG incubation at 50 °C for 2 minutes, AmpliTaq Gold DNA Polymerase activation at 95 °C for 10 minutes, and 40 cycles of denaturation (95 °C for 15 seconds), annealing and extension (60 °C for 1 minute). The specificity of primers was evaluated by visualization of PCR products on agarose gel electrophoresis. PCR efficiency was above 95 %. Relative quantification of gene expression was performed using the standard curve method comprising four serial dilution points (ranging from 0.1 to 50 ng). Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) gene expression was used for data normalization. Additionally, phospho-glycerokinase (PGK) was also used as a control gene for the quantification of *CLN2*/mRNA.

The primer and probe sequences were designed using Primer Express software 2.0 (Applied Biosystems, USA), and are as follows: *CLN2* primer 5'-CAT CCA GCG GGT CAA CAC T-3' and 5'-TGT CTT CCA GAG ACA GAC CAA CA-3'; *CLN2* probe 5'-CAC CCT GCT CTT CGC CTC AGG TGA C-3' located at the exon8/exon9 boundary, or assay-on-demand Hs00166099_m1 (Applied Biosystems, USA) with the probe located at the exon3/exon4 boundary; *CLN3* primer 5'-CGT GAA CAC CTT CCA CAA CAT C-3' and 5'-GGC CGC CAT TGC AAA CT-3'; *CLN3* probe 5'-CTG GAG ACC AGT GAT GAG-3' located at the exon13/exon15 boundary; *CLN5* primer 5'-GCC CCT TTC TGG TGT AAT CAA G-3' and 5'-TCC CAT TTT CCT TCC AGT GAA-3'; *CLN5* probe 5'-CTG CCT GCT TTT TTG AGG GAA TTG ATG A-3' located to exon3. Primers and probes for GAPDH and PGK were ordered as ‘‘assays on demand’’ from Applied Biosystems (USA). For each sample, the RNA extraction procedure was repeated twice, each batch analysed independently three times and the average cycle threshold (C_t) values from duplicate PCRs calculated. All genes exhibited an interassay and intersample C_t variance ≤0.10. The relative quantification of the mRNA was determined through the efficiency-calibrated model (Pfaffl, 2001). The *t* test was used to assess differences in gene expression

between patient and control cells. Two-sided *P* values of inferior to 0.05 were considered statistically significant. Data were expressed as means ± SD (standard deviation).

8.9 Antibodies

Polyclonal antibodies used for the detection of the endogenous CLN2 and CLN5 proteins were raised in rabbit against synthetic peptides corresponding to amino acids 204-218 (VIRKRYNLTSQDVGS) and 105-119 (PYKRFDFRPKPDPYC) (ACP33) of the precursor human CLN2 and CLN5 proteins respectively (Eurogentec, Belgium). The overexpressed CLN5 protein was also detected by a C-terminal peptide specific antibody (C/32) corresponding to amino acids 393-407 (EEIPLPIRNKTLSGL) of the human protein (Eurogentec, Belgium), by a previously described rabbit antibody (1RmI-4) raised against a GST-Cln5 protein (amino acids 40-284) or by a guinea pig antibody (1GmII-3) raised against a GST-Cln5 protein (amino acids 40-341) (Holmberg *et al*., 2004). Several commercially available antibodies were puchased from Santa Cruz Biotechnology (Santa Cruz, USA): rabbit anti-human GAPDH (sc-25778); goat anti-human calregulin (sc-6467); anti-human calnexin (sc-6465); mouse monoclonal anti-GFP antibody. The CLN3 protein was detected by the use of the peptide specific antibody 385 (Luiro, *et al*., 2001). Human (H4A3) and mouse (1D4B) LAMP-1 specific antibodies were obtained from the Developmental Studies Hybridoma Bank, developed under the auspices of the NICHD (National Institutes of Health), and maintained by the University of Iowa (USA). The mouse monoclonal PDI (Protein Disulfide Isomerase) antibody (1D3) was obtained from Stressgen (USA).

8.10 Cloning of cDNA and other plasmid constructs

The pCMV expression vector carrying the human wild type *CLN5* cDNA (aa 1-407) and mutated human *CLN5* with the disease causing mutations p.Y392X (Fin_M), p.E253X (SWE) and p.D279N (EUR), have been described previously (Isosomppi *et al*., 2002; Vesa *et al.*, 2002). Human CLN5-pCMV carrying the p.R112P (POR) and p.R112H (COL) mutations were generated by site-directed *in vitro* mutagenesis using the QuickChange mutagenesis kit (Stratagene, USA), following the manufacturer's instructions. The GFP-CLN5 constructs were generated by cloning M1 to M4 *CLN5* cDNA with an N-terminal GFP-tag in the pEGFP- C1 vector (Clonetech, USA), according to the manufacturer's instructions.

8.11 Cell fractionation

Human skin fibroblasts were trypsinized, washed with cold 0.9 % NaCl, and then centrifuged. The cell pellet was and frozen at -70 °C. Cell pellets from two cultures of growth area of 175 cm^2 at 70–80 % confluence each were re-suspended in lysis buffer (0.25 M sucrose, 5 mM imidazole, and 1 mM, pH 7.4, EDTA, PMSF, and protease inhibitor cocktail from Sigma) and lysis was performed on a glass homogenizer on ice. The extract was centrifuged at 335*g* rpm for 10 min and the supernatant collected was centrifuged at 14200*g* for 15 min. The resulting pellet was re-suspended in lysis buffer and the total protein content estimated by the Bradford protein assay.

8.12 Selective extraction of proteins

Approximately 1 mg of total protein was used for each extraction. Samples were lysed by sonication in 300 µL of either a saline buffer (0.3 M NaCl, 20 mM, pH 8.0, MOPS, 1 mM, pH 8.0, EDTA, 1 mM, pH 8.2, EGTA, PMSF, and a protease inhibitor cocktail from Sigma) or a sodium carbonate 0.11 M solution. After lysis, 900 µL of the respective buffer was added and mixture were incubated for 45 minutes on ice. A small aliquot was taken for posterior analysis and the samples were centrifuged for 1 hour at 45,000 rpm and 4 °C (Sorvall OTD75B). Subsequently, the pellet was re-suspended in Triton X-100 0.1 % and the supernatant was precipitated with trichloroacetic acid (TCA). Protein precipitation by TCA was performed by adding 1/ 9 volumes of TCA to the samples, followed by a 30 minutes incubation on ice. Samples were then centrifuged for 15 minutes at 14200*g* and 4 °C. After removal of the supernatant, 1 mL of acetone was added to the pellet and the mixture was vortexed. After another centrifugation the supernatant was removed and the pellet left to dry. Soluble and insoluble factions obtained in both procedures were re-suspended in gel loading buffer (GLB) and analysed by immunoblotting.

8.13 N-Glycosylation analysis

Cell lysates were subjected to digestion with either Endo H (New England BioLabs, USA) or PNGase F (Roche, Switzerland) enzymes for N-glycosylation analysis. Reactions occurred at 37 ºC for 24 hours. Lysates were prepared by adding 1 µL of 10x Glycoprotein Denaturating Buffer to 20 µg of total protein and water to a final volume of 10 µL. Detailed reaction mixtures, in a final volume of 20 µL, were as follows: Endo H treatment - 2 µL 10x

G5 reaction buffer, 5-7 µL H2O and 1-3 µL Endo H; PNGase F treatment - 2 µL 10x G7 reaction buffer, 2 µL NP-40, 3 µL H2O and 3 µL PNGase F. The reaction was stopped and the proteins denatured by heating at 100 °C for 10 minutes. The products of digestion were analysed by immunoblotting, according to the procedures described in this thesis.

8.14 Detection of endogenous protein by immunoblotting

Cell lysates were prepared according to either the cell fractionation procedure, or the selective extraction protocol, described in this chapter. Approximately 60 µg of total protein for each sample, prepared in gel loading buffer containing dithiothreitol (DTT) and heat denatured, were separated in 13 % acrylamide SDS-PAGE gels on a Mini-V electrophoresis system (Gibco), according to the standard Laemmli method (Laemmli, 1970). Proteins were transferred to a nitrocellulose membrane (Protan-Millipore, USA) by electroblotting and Ponceau protein stain was used to confirm protein loading and transfer. A tris-buffered saline (TBS) solution with 5 % dried non-fat milk was used as blocking buffer and to dilute primary antibodies. Incubation with the primary antibodies occurred overnight at 4 °C. Incubation with the secondary antibody (Pierce, USA), conjugated to horse radish peroxidase (HRP) and diluted in TBS with 0,1 % Tween (TBS-T) was performed for 2 hours at room temperature. The “Super Signal West Dura extended duration substrate kit” from Pierce (USA) was used as HRP substrate for the detection of antigens. Relative quantification of CLN2 and CLN5 proteins was performed by densitometric scanning of the film using either Calnexin or GAPDH as endogenous protein markers, respectively.

8.15 Detection of over-expressed protein by immunoblotting

Transient transfections of Hela, COS-1, and SH-SY5Y cells were performed on 6-well plates with the Fugene HD (Roche, Switzerland) transfection reagent, following the manufacturer’s instructions. Forty-eight hours after transfection the cells were collected and lysed for 20 minutes, on ice, in the following buffer: 50 mM tris-HCl pH 7.5, 300 mM NaCl, 1 % TritonX-100, and protein inhibitor cocktail (Roche, Switzerland). Lysates were centrifuged for 10 minutes at 14200*g* at 4 ºC. Gel loading buffer containing β-mercaptoethanol was added to lysates supernatant and after heat denaturation the samples were loaded and separated on 10-11 % acrylamide SDS-PAGE gels, according to the standard procedure (Laemmli, 1970). Proteins were transferred to nitrocellulose

membranes (Amersham Biosciences, UK) by electroblotting and Ponceau protein stain was used to confirm protein loading and transfer. A tris-buffered saline (TBS) solution with 0,1 % Tween and 5 % dried non-fat milk was used as blocking buffer and to dilute the primary and secondary antibodies. Incubation with primary antibodies and secondary HRP-conjugated antibodies (DAKO, Denmark) was performed at room temperature for 2 hours. The ECL Western Blotting Detection Kit (Amersham) was used as HRP substrate for the detection of antigens.

8.16 Microscopy and protein detection by immunofluorescence

Hela, COS-1, and SH-SY5Y cells were cultured on cover slips in 6-well plates and transiently transfected with the Fugene HD (Roche, Switzerland) transfection reagent, following the manufacturer's instructions. Approximatively 48 hours after transfection cells were fixed with ice cold methanol for 3-5 minutes. The cover slips were incubated with the primary antibodies, diluted in 0.5 % BSA in PBS, for one hour at 37 ºC. Incubation with the secondary antibodies, diluted in 0.5 % BSA in PBS, was performed at room temperature for one hour. Secondary antibodies were acquired from Jackson ImmunoResearch (USA). The labelled coverslips were mounted using GelMount (Biomeda Corp., USA) and visualized using a Leica DMR confocal microscope with TCS NT software (Leica Microscope and Scientific Instruments Group). ImageJ, Adobe Photoshop and Adobe Illustrator were used for image processing and figure assembly. Each channel was adjusted independently for brightness and contrast.

8.17 Software prediction tools

Online servers for GENSCAN (http://genes.mit.edu/GENSCAN.html) and SpliceView (http://bioinfo.itb.cnr.it/oriel/splice-view.html) were used to predict the locations of splice sites and the exon-intron structures of the *CLN2* gene. The strengths of 5' and 3' splice junctions in the *CLN2* gene were estimated with the MaxENT model (http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq_acc.html). For 5' and 3' splice sites a log-odd ratio (MaxENT score) is assigned to a 9-mer or a 23-mer, respectively. The higher the score, the higher the probability that the sequence is a true splice site. Ideal MaxENT splice site scores are: 11.81 (5'-ss) and 13.59 (3'-ss). The web tools ESEfinder (http://rulai.cshl.edu/tools/ESE/) and RESCUE-ESE (http://genes.mit.edu/burgelab/rescue-ese/) were used to predict and identify possible

exonic splicing enhancers.

Two models were used to predict the presence and location of signal peptides and signal peptide cleavage sites in the human CLN5 protein: SignalP and SigCleave. SignalP (http://www.cbs.dtu.dk/services/SignalP), uses two different neural networks, one for predicting the actual signal peptide and one for predicting the position of the signal peptidase I (SPase I) cleavage site. SigCleave (http://emboss.bioinformatics.nl/cgi-bin/emboss/sigcleave) predicts the site of cleavage between a signal sequence and the mature exported protein using the method of von Heijne. The scoring weight value for the predicted cleavage site should be at least 3.5.

9 Results

The most relevant results obtained in the course of the PhD project were submitted to international publications for independent peer-review. Two papers were accepted for publication and a draft-manuscript for a third paper has been submitted for review (appendix). The intent of this chapter of the thesis is to provide an overall and integrated perspective of the achieved results. The localization of the described results with regard to the mentioned publications are as follows:

9.1 LINCL: evidence for mild disease due to an intron retention defect

A Portuguese patient exhibiting a juvenile NCL clinical spectra and curvilinear inclusions upon ultrastructure examination was investigated for the genetic variant. The age of onset of symptoms was at 10 years and the patient presented a progressive cognitive and motor degeneration, epilepsy and seizures. Preliminary *in vitro* enzymatic assays using synthetic substrates revealed a deficient activity of TPP1/CLN2 in patient fibroblasts and leukocytes cell extracts, establishing the genetic variant as CLN2.

9.1.1 Genotype identification

The genomic DNA was screened for mutations in the *CLN2* gene and a novel alteration, g.4196A>G, was found in homozigosity (Figure 7). Intronic sequence based primers were used and screening of exons and flanking intron regions revealed no other mutations. As expected, both parents were found to be heterozygous for the same gene defect. This alteration was not encountered in 100 ethnically matched control chromosomes.

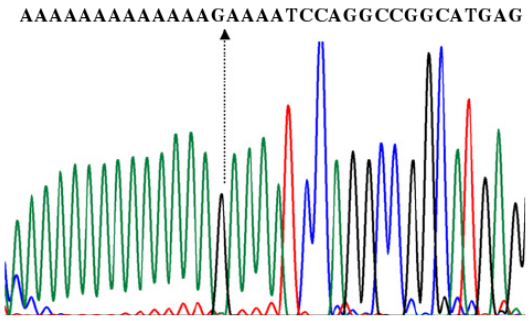


Figure 7. Electropherogram showing the genomic nucleotide sequence across the mutation site in the intron 7 3' region of the *CLN2* gene. The arrow indicates the mutation site.

The identified mutation occurs in a non-coding region, near the wild type acceptor splice site (ss) in intron 7. Interestingly, it is predicted to create a new possible acceptor

splice site which, if used by the cell, would result in the retention of nine nucleotides in the mRNA molecule and, consequently, in the in-frame insertion of three amino acids between proline295 and glycine296. In order to evaluate this possibility, and to determine the strength of the newly created acceptor site, several splice site prediction software models were used to analyse and compare the mutated and wild-type gene sequences.

9.1.2 Predictions from *in silico* analysis

The possibility that the mutation g.4196A>G creates a novel acceptor splice site led us to explore the strength of donor and acceptor splice sites, intronic/exonic splicing enhancers and silencers, and protein binding sites in the *CLN2* gene.

The most relevant data was obtained with the MaxEntscan model (Yeo and Burge, 2004). MaxEntscan accesses the strength of donor and acceptor splice sites in gene sequences. Low and negative scores are associated with increased susceptibility to splice mutations and alternative splicing, thus suggesting a weak splice site, while higher scores point to a more efficient splicing. MaxEntscan attributed a negative score of -1.08 to the wild type 3'-ss present in intron 7 (Figure 8). On the other hand, the calculated score for all other 3' splice sites present in the *CLN2* gene is positive and averages a value of 7.239.

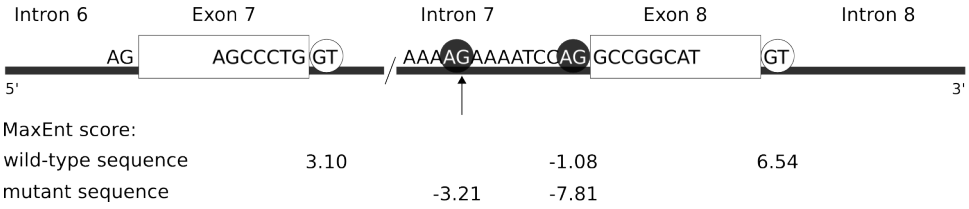


Figure 8. Schematic representation of intron 7 of the CLN2 gene and the scores obtained with the MaxEntscan model for the strength of the donor and acceptor splice site in the vicinity of mutation g.4196A>G. Position of the mutation is indicated by an arrow. Donor splice sites are surrounded by a white circle, while the wild-type and novel acceptor 3'-ss are inside a black circle.

Analysis of the mutated sequence also revealed a negative score for the novel splice site (-3.21). Moreover, in the presence of the novel 3' ss, the wild-type splice site is

attributed a lower score (-7.81). In theory, despite being negative, the higher score of the novel splice site suggests that it could be preferably used by the cell. However, the wild-type sequence is still present and theoretically active and so could compete with the newly created splice site. Other critical factors for splicing mechanisms such as intronic/exonic splicing enhancers (ISE/ESE) and silencers (ISS/ESS) were also evaluated. Looking for serine/arginine-rich (SR) protein binding sites using the ESEFinder (Cartegni *et al*., 2003) software yielded no relevant additional information. Nevertheless, the Rescue-ESE (Fairbrother *et al*., 2002) prediction model suggested four potential splice site enhancer sequences downstream of the new 3'-ss.

The identified IVS7-10A>G mutation had not been reported before and so several studies were conducted to analyse the impact of the alteration at the molecular level.

9.1.3 Molecular consequences of the IVS7-10A>G mutation

In order to understand how the identified mutation affects protein synthesis and function, several approaches were delineated which included analysis of mRNA, protein maturation and enzyme activity.

The *CLN2* mRNA was studied through Northern blot and quantitative real time RT-PCR (qRT-PCR). Northern blot analysis revealed two bands co-migrating with the control RNAs and, under the experimental conditions used, no other mRNA specie was observed (Figure 9a). Cells from a classic R208X homozygous patient were used as control and as expected no bands were detected. The level of *CLN2* mRNA was quantified by qRT-PCR, which showed a normal steady-state level of *CLN2* mRNA in the studied patient (Figure 9b). These data suggests IVS7-10A>G mutation does not cause alternative splicing or interferes with mRNA stability or degradation.

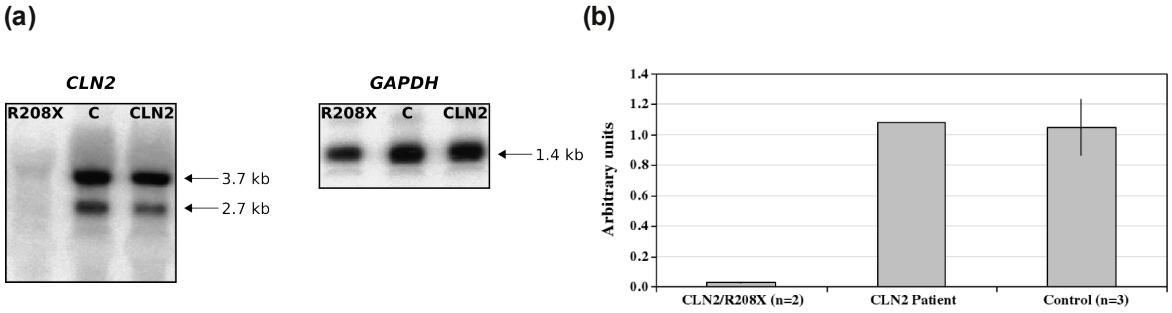


Figure 9. (a) Northern blot analysis of *CLN2* mRNA from cultured human fibroblasts. *GAPDH* cDNA was used as control probe. C, Healthy control individual; R208X, LINCL/CLN2 patient homozygous for the nonsense mutation p.R208X; CLN2, Patient under study. (b) Relative *CLN2* mRNA quantification by real-time PCR using the mean value of the expression of *GAPDH* and *PGK* for data normalization. Values were calculated in reference to the normal controls whose expression was arbitrarily attributed the value of 1.0. Dr. Carla Teixeira is acknowledged for the figures and her assistance with the Northern blot and qRT-PCR analysis.

To address the question of which splice site, wild-type or novel, is in fact used by the cell, the *CLN2* cDNA was sequenced. Data revealed the presence of only one transcript with 9 additional nucleotides, derived from the 3' extremity of intron 7, when compared to the wild-type molecule. This observation is consistent with the usage of the newly created acceptor splice site (Figure 10).

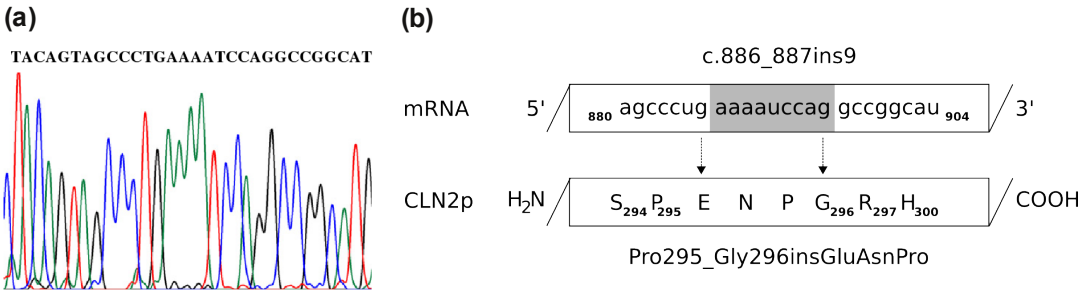


Figure 10. (a) Electropherogram showing the insertion of the last 9-nt of intron 7 in the *CLN2* cDNA from the patient. (b) Schematic diagram depicting the predicted effect of the mutation on the mature *CLN2* mRNA and primary protein structure.

Additionally, RT-PCR products were also analysed by agarose gel electrophoresis (Figure 11a). Only one band was observed for the patient sample, with a slower migration when compared to the control. Quick screening for the IVS7-10A>G mutation was established via SSCP and heteroduplex analysis (Figure 11b). The pattern observed for the heteroduplexes were also in agreement with the presence of only one, longer, transcript.

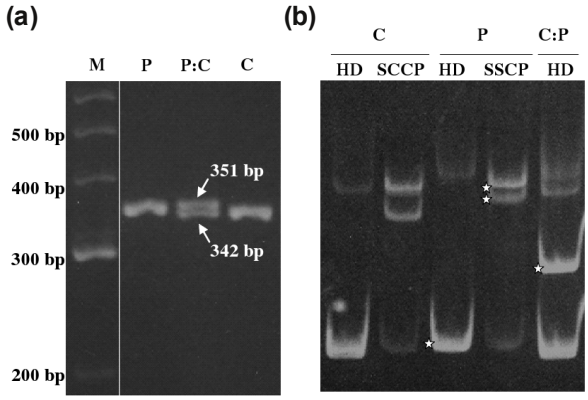


Fig. 11. Agarose gel electrophoresis of RT-PCR products generated by amplification of exons 6–8 of the CLN2 gene. (a) The products of RT-PCR were analysed on ethidium bromide-stained non-denaturing polyacrylamide gel after heat denaturing and cooling in ice water bath (single strand analysis) or room temperature (heteroduplex analysis) (b). HD, heteroduplex; SSCP, single strand conformational polymorphism; M, molecular-weight markers; C, healthy control individual; P, CLN2 patient, C:P, mix of RT-PCR products from control and CLN2 patient (1:1); *, mobility shift.

Overall, the results showed that only the newly created acceptor slice site in intron 7 is being used, resulting in the production of an alternative transcript. Moreover, no evidence for alternative splicing was found. RT-PCR is a high sensitivity technique and so the possibility that this process may be occurring at an undetectable level is unlikely.

Translation from the mutated sequence results in a predicted polypeptide with three extra amino acids after proline 295 (p.Pro295_Gly296insGluAsnPro) and so the impact of the IVS7-10A>G mutation on the maturation and level of the CLN2 protein was analysed by Western blotting. Endogenous protein was detected with the use of a novel peptide specific anti-human TPP1 polyclonal antibody. Specificity of the antibody was accessed by comparison with a p.R208X homozygous CLN2 patient as this mutation results in the complete absence of protein. Densitometric analysis of the blot showed a reduction of

approximately 40 % in the level mutant TPP1 (Figure 12), suggesting that the mutation may interfere with the stability of the protein. Although the insertion of 3 amino acids in the CLN2 polypeptide could result in altered post-translation processing, no discernible difference in molecular weight was observed when compared to the wild-type protein.

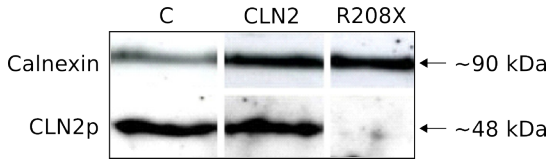


Figure 12. Western blot analysis of CLN2 protein in control and CLN2-deficient fibroblasts. The antibody was produced against a synthetic peptide encompassing amino acids 204-218 of the precursor human CLN2 protein. Calnexin was used as endogenous control. C, Healthy control individual; R208X, LINCL/CLN2 patient homozygous for the nonsense mutation p.R208X; CLN2, Patient under study.

Since no dramatic consequences of the mutation were observed for both protein stability and maturation, the enzymatic activity was carefully analysed. The activity of TPP1 was determined in both leukocytes and fibroblasts, using a synthetic substrate, and was found to be markedly deficient in both cell types. The patient exhibited approximately 4 % and 2 % of control TPP1 activity in leukocytes and fibroblasts respectively (Table 2). Moreover, this activity was about two fold higher when compared with the classical p.R208X homozygous CLN2 patient.

Table 2. TTP1 activity in leukocytes and fibroblasts

	Enzymatic activity		
	PPT1	TPP1	ß-Gal
Leukocytes			
LINCL/CLN2/R208X (n=2)	87	4.6	357
CLN2 Patient	23	9.9	244
Controls (n=25)	63(±23)	251(±56)	284(±85)
Fibroblasts			
LINCL/CLN2/R208X (n=2)	146	5.9	512
CLN2 Patient	119	11	415
Controls (n=10)	152(±35)	482(±113)	813(±280)

The enzymatic activity was expressed in nmol/h/mg protein. Data is reported as mean activity of all cases studied, each analysed independently at least twice. PPT1, Palmitoyl protein thioesterase; TTP1, Tripeptidyl peptidase; ß-Gal, ß-galactosidase used as reference enzyme in the assays; Dr. Carla Teixeira is acknowledged for the data and the enzymatic activity assays.

9.2 Insights into molecular mechanisms of CLN5 deficiency

The Finnish vLINCL subtype is specially prevalent in the Finnish population and at the time this project was initiated only two cases had been reported outside Northern Europe. This study began with the identification of the first Portuguese patient with $vLINCL_{Fin}$.

Progressive attention deficits and regression of language capacities were noticed at 3 years of age and motor clumsiness appeared one year later. At the age of 7 years old visual deterioration started to develop, accompanied by gait ataxia, myoclonus and generalized epileptic seizures. Electron microscopy analysis of secretory sweat glands cells revealed a mix of intracellular inclusions consisting of a mix of curvilinear and fingerprint profiles which, combined with the described age of onset and clinical spectra, was suggestive of a variant late-infantile phenotype. Supporting this hypothesis, CLN1/PPT1 and CLN2/TPP1 enzyme activity were evaluated and found to be normal (data not shown).

9.2.1 Identification of two novel *CLN5* mutations underlying $vLINCL_{Fin}$

Aiming to identify the patient's genotype, mutational screening was performed for *CLN3*, *CLN5* and *CLN6* genes. No sequence variations were observed for the coding sequences and intronic flanking regions of the *CLN3* and *CLN6* genes. However, 3 mutations were identified in the *CLN5* gene. The patient was found to be a compound heterozygous: in the maternal allele, a novel c.335G>C (p.R112P) mutation and the previously described c.835G>A (p.D279N) (Savukoski *et al*., 1998) mutation were found, while in the other allele a novel nonsense (c.565C>T/p.Q189X) mutation was discovered (Figure 13).

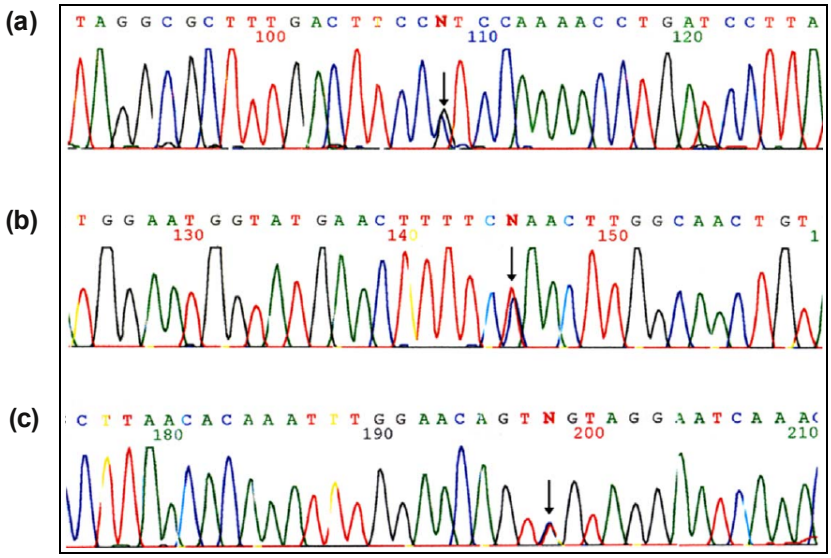


Figure 13. Electropherograms showing genomic nucleotide sequence across the mutation site. (a) Missense mutation c.335G>C (p.R112P) in exon 2 (sense direction); (b) Nonsense mutation c.565C>T (p.Q189X) in exon3 (sense direction); (c) Missense mutation c.835G>A (p.D279N) in exon 4 (anti sense direction); Mutations are indicated by an arrow.

Both missense mutations were not observed in 50 control individuals, partially discarding the possibility of representing common polymorphisms. Moreover, they occur in evolutionarily conserved residues.

9.2.2 Impact of mutations on the mRNA and protein levels

To study the biological impact of the mutations identified in the CLN5 patient analysis at the level of mRNA and protein was then performed. Messenger RNA was quantified by qRT-PCR and the data revealed a 55% decrease in the level of *CLN5*/mRNA (Figure 14) when compared with the control sample. The observed decrease in the level of mRNA is consistent with the hypothesis that the p.Q189X mutation is indeed a null allele, and so the detected mRNA should originate from the allele carrying the two missense mutations.

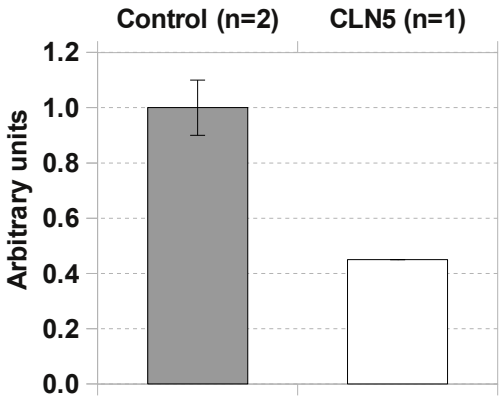


Figure 14. Relative *CLN5*/mRNA quantification by real time PCR. Values were calculated in reference to normal controls whose expression was arbitrarily attributed the value of 1.0. Dr. Carla Teixeira is acknowledged for the figure and the qRT-PCR analysis.

Protein was also analysed by SDS-PAGE followed by Western Blot and immunodetection with the novel ACP33 antibody. This is a rabbit antibody raised against a synthetic peptide corresponding to amino acids 105-119 of the human CLN5 protein. Antibody specificity was assessed by the means of a peptide competition assay (Figure 15a). Moreover, no significant homologies were found between the peptide sequence used for the antibody production and any other known protein sequence. Western blot analysis revealed the presence of 4 bands (49-44 kDa) in organellar enriched lysates from control human skin fibroblasts (HSF). Occasionally only three bands were visible. When compared to the control HSF, the Portuguese CLN5 patient showed a decrease of approximately 50 % in overall protein level (Figure 15b). This observation is in agreement with the data from the mRNA quantification. Additionally, a low molecular weight polypeptide was not detected and so it is unlikely the p.Q189X carrying allele is being translated. Detection of endogenous CLN5 polypeptides has not been reported elsewhere so far.

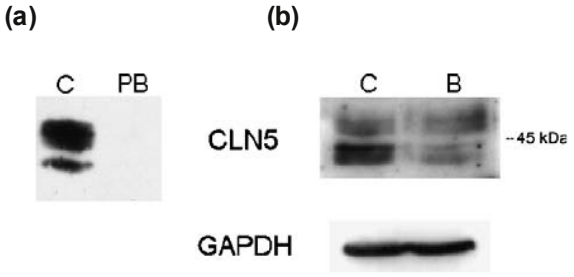


Figure 15. (a) Antibody specificity was assessed by incubation with the peptide used for immunization (PB). (b) Western blot analysis of CLN5 protein in control and $vLINCL_{Fin}$ fibroblasts. CLN5 was detected with the ACP33 antibody. GAPDH was used as an endogenous control for relative protein quantification; C, Normal control; B, $vLINCL_{Fin}$/CLN5 patient.

Initial structural predictions suggested that human CLN5 to be an integral membrane protein, with 2 putative transmembrane domains (Savukoski *et al*., 1998). However, in recent years data supporting the existence of both membrane and soluble CLN5 polypeptides has been presented (Holmberg *et al*., 2004; Vesa *et al*., 2002). We tried to address this issue by studying the molecular forms of CLN5 detected with the ACP33 antibody. Cell lysates were subjected selective protein extraction with either a 0,3 M NaCl or a 0,11 M sodium carbonate buffer.

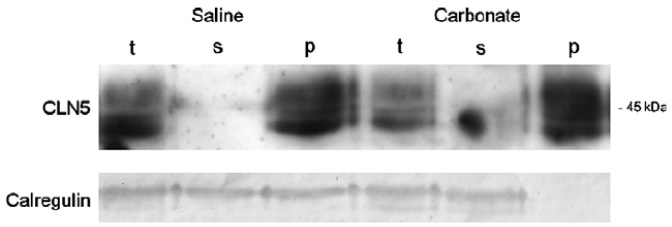


Figure 16. Western blot analysis of CLN5 protein after selective extraction of control fibroblasts protein. For validation purpose analysis of an extrinsic membrane protein was performed using anti-human calregulin antibody; t, Pre-centrifuge sample; s, Supernatant; p, Pellet.

Western blot analysis revealed the presence of all detected polypeptides uniquely in the pellet fraction of both saline and carbonate extraction (Figure 16). As the carbonate extraction only solubilizes peripheral membrane associated proteins this observation suggests that all detected polypeptides are integral membrane proteins. Calreticulin, an ER resident soluble protein, which also occurs as a membrane associated form (Michalak *et al*., 1991), was used here as control, and, as expected, is absent from the pellet of the carbonate extraction.

NCL proteins have for a long time been speculated to intervene in a common unknown biochemical pathway. Co-immunoprecipitation experiments reported by Vesa and colleagues have suggested protein interactions between the CLN5 protein and CLN3/CLN2 (Vesa *et al*., 2002). Interactions at the protein level raises the possibility of transcriptional regulation in the corresponding protein-encoding genes and such a hypothesis was evaluated by searching for gene expression co-regulation at the level of mRNA. The level *CLN2*, *CLN3* and *CLN5* mRNAs was thus analyzed in CLN2, CLN3 and CLN5 diseased fibroblasts (Figure 17).

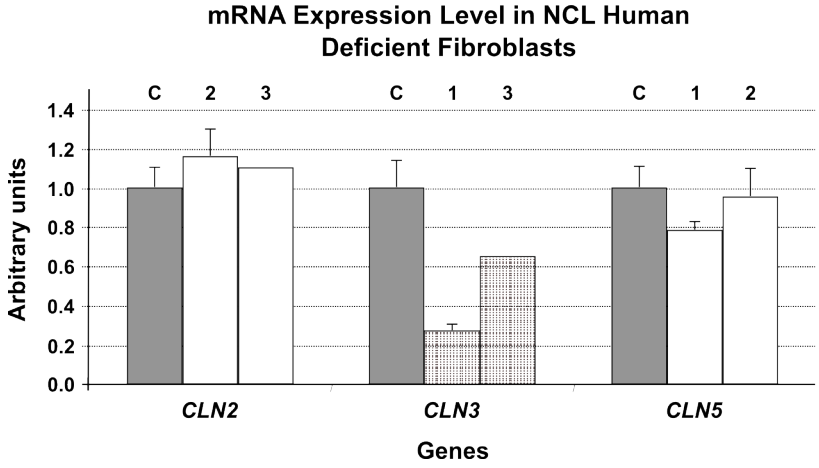


Figure 17. Differential expression of *CLN2*, *CLN3*, and *CLN5* genes in human fibroblasts from patients with CLN2, CLN3, and CLN5 variants. Values were calculated in reference to the control whose expression was arbitrarily attributed the value 1.0. Data represent means ± SD. C—Control fibroblasts from normal individuals; 1—CLN2 deficient human fibroblasts; 2—CLN3 deficient human fibroblasts; 3—CLN5 deficient human fibroblasts; dotted bars correspond to changes with statistical significance. Dr. Carla Teixeira is acknowledged for the figure and the qRT-PCR analysis.

No significant changes were detected for CLN2 encoding mRNA in CLN3 and CLN5 deficient cells when compared to normal control cells. The *CLN3*/mRNA however, was found to be significantly decreased in both CLN2 and CLN5 deficient fibroblasts to approximatively 27 % and 65 % of mRNA levels found in control cells respectively. Moreover, the *CLN5/mRNA* was slightly decreased in CLN2 deficient cells (Figure 16), representing approximately 79 % of the amount detected in control cells.

9.3 Novel insights into the biological properties of CLN5

Attempts to further characterize the CLN5 protein using the ACP33 antibody were pursued. However, over time the ACP33 antibody ceased to function in Western blot applications. This was likely due to suboptimal storage conditions or perhaps excessive thaw/freeze procedures, which eventually lead to the degradation of the antibody. Other CLN5 peptide specific antibodies were produced and tested, but failed to detect any endogenous protein in both Western blotting and immunofluorescence analysis. It is known that cells from NCL patients represent a natural cell model recurrently used not only to explore the molecular pathophysiological mechanisms but also to understand more fundamental aspects of the cell biology. However, the above mentioned constraints led us to extend the study to genetically engineered cultured cells. A partnership was established with Professor Anu Jalanku's NCL research group in Helsinki, Finland, to further study the CLN5 protein. The fruits of this collaboration are presented bellow.

9.3.1 Analysis of N-terminal processing

Several key points of the CLN5 protein biology have yet to be comprehensively characterized. The existence of four initiation AUG codons (M1-4) in the CLN5 mRNA molecule for instance is an aspect that warrants further investigation. It may allow for a mechanism of alternative methionine utilization, which, by creating polypeptide diversity, may bear functional consequences.

Initial theoretical predictions postulated the absence of an ER signal peptide in the CLN5 sequence. However, only the human sequence starting from the first methionine (M1) was considered for analysis (Savukoski *et al*., 1998). Of notice, the human CLN5/M4 sequence shares great homology to the mouse Cln5 protein, which has been described as being soluble (Holmberg *et al*., 2004). We have re-run the simulations for all M1-4 human sequences, using several signal prediction models to evaluate CLN5 sequences starting from all four initiation methionines. Results for the SignalP and SigCleave models are presented in Table 3.

Table 3. Signal peptide prediction for M1-4 CLN5 sequences.

Program		hCLN5 M1	hCLN5 M2	hCLN5 M3	hCLN5 M4	mCln5
SignalP	Signal peptide	No	Yes	Yes	Yes	Yes
	Probability	0.143	0.785	0.987	1	0.999
	aa position	-	62 or 66	46	34	33
	aa pos. (M1)	-	92 or 96	96	96	-
SigCleave	Score	9.11/7.55	9.11/7.55	9.11/7.55	9.4/7.5	8.78/7.75
	aa pos.	89/91	60/62	40/42	28/30	27/28
	aa pos. (M1)	89/91	90/92	90/92	90/92	-

The prediction of signal peptides and the correspondent cleavage sites in human CLN5 sequences (M1-M4) and mouse Cln5 are shown in this table. The position of the predicted signal peptidase I cleavage site is represented in reference to both the initiation methionine used in the analysed sequence (aa position) but also to the sequence starting with M1 (aa pos. (M1)). Additionally the probability of a signal peptide and the overall prediction verdict is also presented for the SignalP model. For the SigCleave prediction only the two highest scores are presented.

The SignalP model (Nielsen *et al*., 1997; Bendtsen *et al*., 2004) predicts the presence of a signal peptide for M2-M4 sequences, although with diminishing likelihoods for the more N-terminal methionines. Since the signal peptide region and predicted cleavage site are common between all M1-4 sequences it is possible that the increasing distance to the N-terminal extremity might lead to these results. SigCleave is another model that predicts the existence of signal peptides and cleavage sites in proteins (von Heijne, 1986; 1987). However, it does not make use of a cut-off to disregard “internal” sites and the distance to the N-terminal extremity. Its predictions are consistent with the previous hypothesis since it proposes identical scores for all M1-4 sequences. Thus, if signal peptide cleavage occurs as these models predict, it is possible that regardless of the initiation methionine used for translation similar polypeptides are produced.

To analyse the issue of alternative methionine initiation and whether it could give rise to distinct polypeptides, we transiently transfected COS-1 and HeLa cells with several constructs created for this purpose.

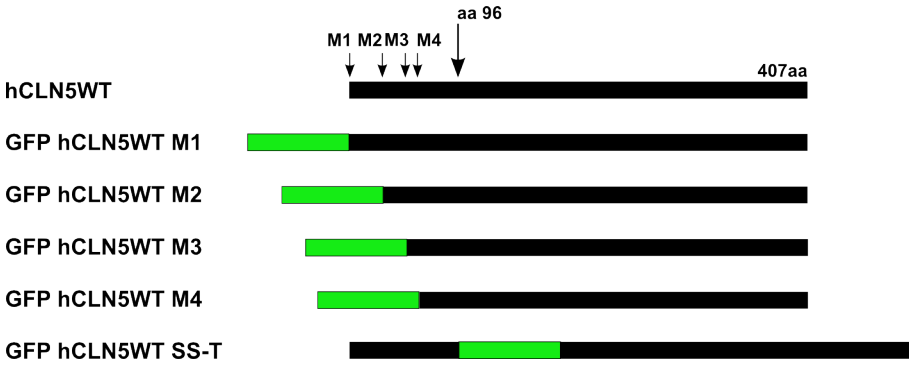


Figure 18. Processing of CLN5. The human CLN5 protein sequence possesses four possible initiation methionines (M1-4) at the amino acid positions of 1, 30, 50 and 62, and one predicted signal peptide cleavage site. An N-terminal GFP-tag was added at the 5' of all possible starting methionines M1-4. These constructs were then used to transiently transfect HeLa and COS-1 cells.

As depicted in Figure 18 a GFP tag was inserted at the 5' of each initiation methionine of the human CLN5 sequence. Additionally, a construct with a GFP tag immediately after the residue 96, between the predicted cleavage site and the recognition site for SPase I, was produced (SS-T). The polypeptides resulting from the overexpression were analysed both by immunoblotting and immunofluorescence. The use of GFP and a novel human specific C-terminal CLN5 antibody (C/32) allowed for the detection of both extremities of the polypeptides. The specificity of the C/32 antibody was assessed by comparison with samples from the non-transfected cells.

Western blot analysis showed GFP and CLN5 specific antibodies detected distinct polypeptides (Figure 19a-b). The C/32 antibody recognized a single 60kDa polypeptide for all M1-4 transfected constructs. On the other hand, the GFP antibody detected several polypeptides with variable molecular weights (43, 41, 39,5 and 38 kDa), which correlate well with the estimated MW for the N-terminal fragment assuming a signal peptide cleavage at aa 96 of the M1 sequence. The lower molecular weight polypeptides detected by the GFP antibody were variable between experiments and, as such, are likely to derived from protein degradation.

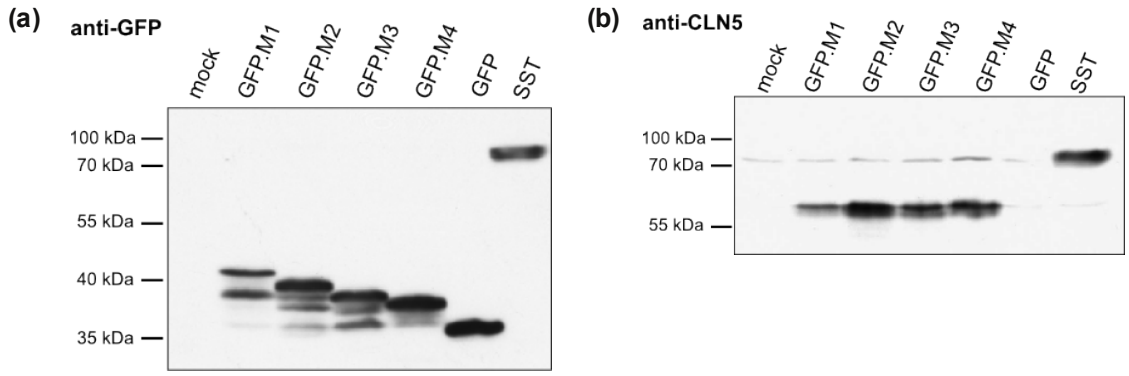


Figure 19. Proteolytic cleavage of the GFP-(M1-M4) CLN5 proteins in COS-1 cells was verified by analyzing cell lysates through SDS-PAGE and immunoblotting with an anti-GFP antibody (a) and with a C-terminal CLN5 specific antibody (b).

Data obtained for the indicated SST construct suggests that the region nearby residue 96 is in fact important for the signal peptide removal. As it can be observed in Figure 19 (a,b), the insertion of the GFP tag at position 96 has interfered with the cleavage process as the same polypeptide was detected by both antibodies. Overall, the data seems to suggest that regardless of the initiation methionine used by the cell only one polypeptide is produced. There is, however, the possibility that other isoforms are being simultaneously produced from other methionines but are not being detected by the used antibody.

Immunofluorescence analysis with the C-terminal antibody revealed the localization of the over-expressed polypeptides (Figure 20). In agreement with what was observed by Western blotting, GFP and the C/32 antibody do not co-localize, indicating that the N-terminal and C-terminal regions of the protein are separated.

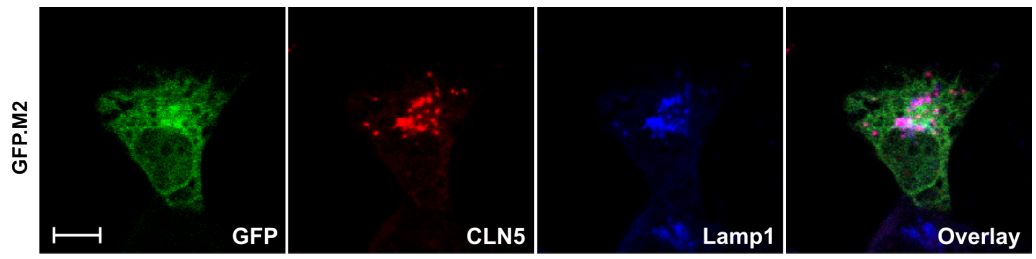


Figure 20. Proteolytic cleavage of the GFP-(M1-M4) CLN5 proteins in COS-1 cells was also verified by comparing the CLN5-specific staining pattern to that of the GFP by using a confocal microscope. LAMP-1 staining was used to indicate lysosomes (GFP.M2 shown as an example). Co-localization between Lamp-1 and CLN5 is indicated by pink. Scale bar - 10 µm.

The C-terminal CLN5 specific antibody presented a vesicular staining and was found to co-localize with the Lamp-1 antibody, pointing to a lysosomal localization. The GFP fluorescence distribution, perinuclear and reticular, resembles an ER-like pattern. The lysosomal localization of the CLN5 polypeptides was also confirmed by detection with other CLN5 specific antibodies, 1RmI-4 and 1GmII-3, described in the “Material and methods” chapter (data not shown). In the case of the SS-T construct, the GFP and CLN5 antibody fluorescence co-localize and present a distribution pattern very similar to the Golgi-complex (data not shown). In a subpopulation of cells, a very strong ER-like staining was observed for the CLN5 protein. However, this is likely an artifact resulting from the transient transfection procedure, which may lead to a cell overload with overexpressed protein.

9.4 *Impact of mutations in the cell biology of CLN5*

9.4.1 Study of the intracellular trafficking and maturation of the CLN5 protein

Previously described mutations in the *CLN5* gene have not been comprehensively documented in preceding publications. Thus, the impact of the most relevant reported mutations to date, including the recently identified p.R112H and p.R112P, in the intracellular trafficking and maturation of the CLN5 protein was studied.

HeLa cells were transfected with constructs based on the human wild-type *CLN5* but carrying the following mutations: p.R112P/POR, p.R112H/COL, p.Y392X/Fin_M, p.D279N/EUR and p.E263X/SWE; the localization of the mutant polypeptides was determined through immunofluorescence using CLN5 specific antibodies. Data showed a different intracellular localization for all polypeptides when compared to the wild-type protein. All analysed overexpressed mutants localized to the ER in HeLa cells, co-localizing well the ER marker protein PDI (Figure 21). Additional incubation of cells with

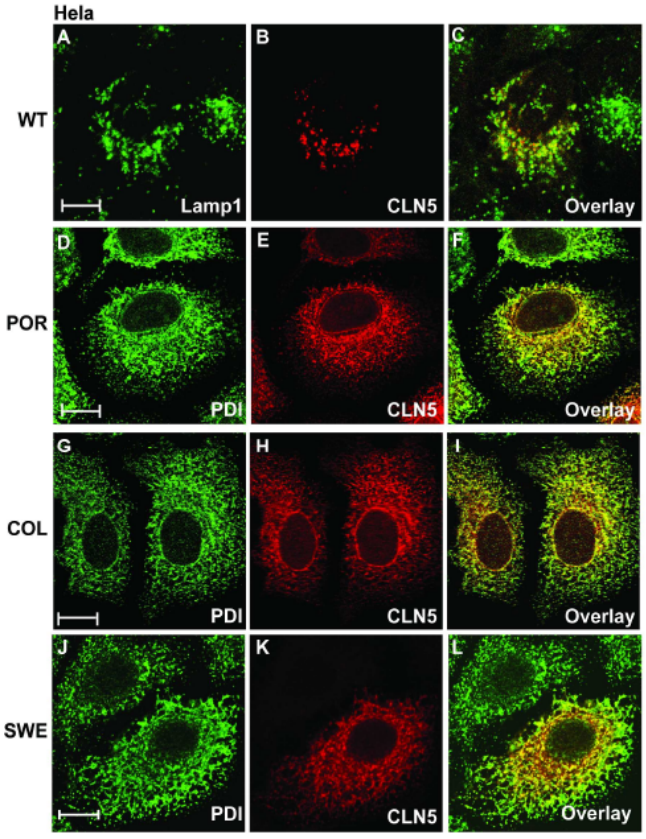


Figure 21. Steady state localization of WT and mutated CLN5 in HeLa cells. Wt CLN5 (A-C) and CLN5 constructs carrying the mutation p.R112P (POR) (D-F), p.R112H (COL) (G-I) or p.E253X (SWE) (J-L) were transiently transfected in HeLa cells. Localisation of the different CLN5 proteins was studied by immunofluorescence labelling using a CLN5 specific antibody (B, E, H, K). PDI was used to label the ER (D, G, J) and LAMP-1 was used to detect lysosomes (A). Co-localisation is indicated in yellow. Scale bar - 10 µm.

CHX, followed by a 5 hour chase, revealed that the p.Y392X and p.E263X were not detectable after 4 hours. Moreover, the EUR mutant polypeptide presented a lysosomal localization in approximately 10% of cells (data not shown).

The mutant p.R112H, p.R112P and p.D279N polypeptides were further analysed by Western blotting. Their N-glycosylation status was studied by digestion of cell lysates with EndoH or with PNGaseF enzymes, which confirmed the presence of both high-mannose residues and complex-type oligosaccharides (Figure 22). In addition, a higher MW polypeptide was observed for the p.D279N mutant, with approximately 65 kDa, suggesting that the new consensus sequence for N-glycosylation created by the p.D279N mutation is in fact used *in vivo*.

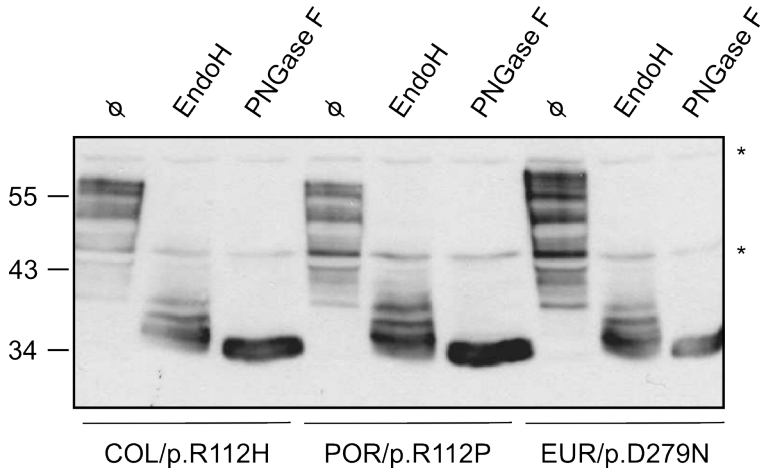


Figure 22. N-Glycosylation analysis of CLN5 polypeptides carrying vLINCL disease causing missense mutations. COS-1 cells were transiently transfected with CLN5 constructs carrying the missense mutations p.R112H (COL), p.R112P (POR) or p.D279N (EUR) for 48 h. Cell lysates were digested with either EndoH or PNGaseF. Untreated (Φ) and treated samples were immunoblotted and CLN5 was detected with C-terminal human CLN5 antibody (C/32). Unspecific bands are indicated with an asterisk (*). Dr. Mia-Lisa Schmiedt is acknowledged for the figure and the N-glycosylation analysis.

In light of these data, the three missense mutations (p.R112P, p.R112H and p.D279N) do not appear to significantly impact the stability and maturation of the protein. Moreover, the localization of these mutant polypeptides to the ER was unexpected since the addition of complex-type oligosaccharides occurs mainly in the *trans* region of the Golgi complex. Such observation suggests a yet uncharacterized retrogade transport

event from the Golgi to the ER could explain this intracellular localization for fully glycosylated mutant CLN5 polypeptides.

9.4.2 NCL-NCL protein interactions

Interaction of CLN5 with the CLN2 and CLN3 proteins have been previously suggested (Vesa *et al*., 2002). Moreover, data has been reported that suggest CLN5 interacts with several NCL proteins, including CLN1. It has also been observed that co-expression of CLN1/PPT1, and not other NCL proteins, with the CLN5 disease causing mutation p.Y392X/FIN_M, resulted in the aparent correction of the trafficking of the mutant protein to the lysosomes (Lyly *et al*., *in press*). Based on this observation the study of NCL protein interaction and $vLINCL_{Fin}$ causing mutations was further explored to evaluating whether PPT1 could also facilitate the trafficking of the p.R112P/POR, p.R112H/COL CLN5 mutants described previously. Although these results were not included in Publication-III, they are described in this section of the thesis as unpublished data.

SH-SY5Y cells were transiently transfected with both wild-type PPT1 and human CLN5 cDNA carrying the mutations: p.R112P/POR, p.R112H/COL and p.Y392X/Fin_M. Localization of mutant protein was analysed by immunofluorescence under a confocal microscope, using Lamp-1 as an endosomal/lysosomal marker. We have shown before that single transfection of these mutated CLN5 constructs results in ER retention of the polypeptides in HeLa cells (Figure 21). Similar observations were made for SH-SY5Y cells (Figure 23). Mutants exhibit mainly an ER resembling pattern, with some ocasional minor lysosomal localization. However, it was observed that co-transfection with wild type PPT1 altered the localisation of all analysed mutant CLN5 proteins. When PPT1 was simultaneously co-expressed, mutant CLN5 exhibited a perinuclear vesicular distribution, co-localising with both PPT1 and Lamp-1 antibodies, pointing to a lysosomal localization (Figure 23).

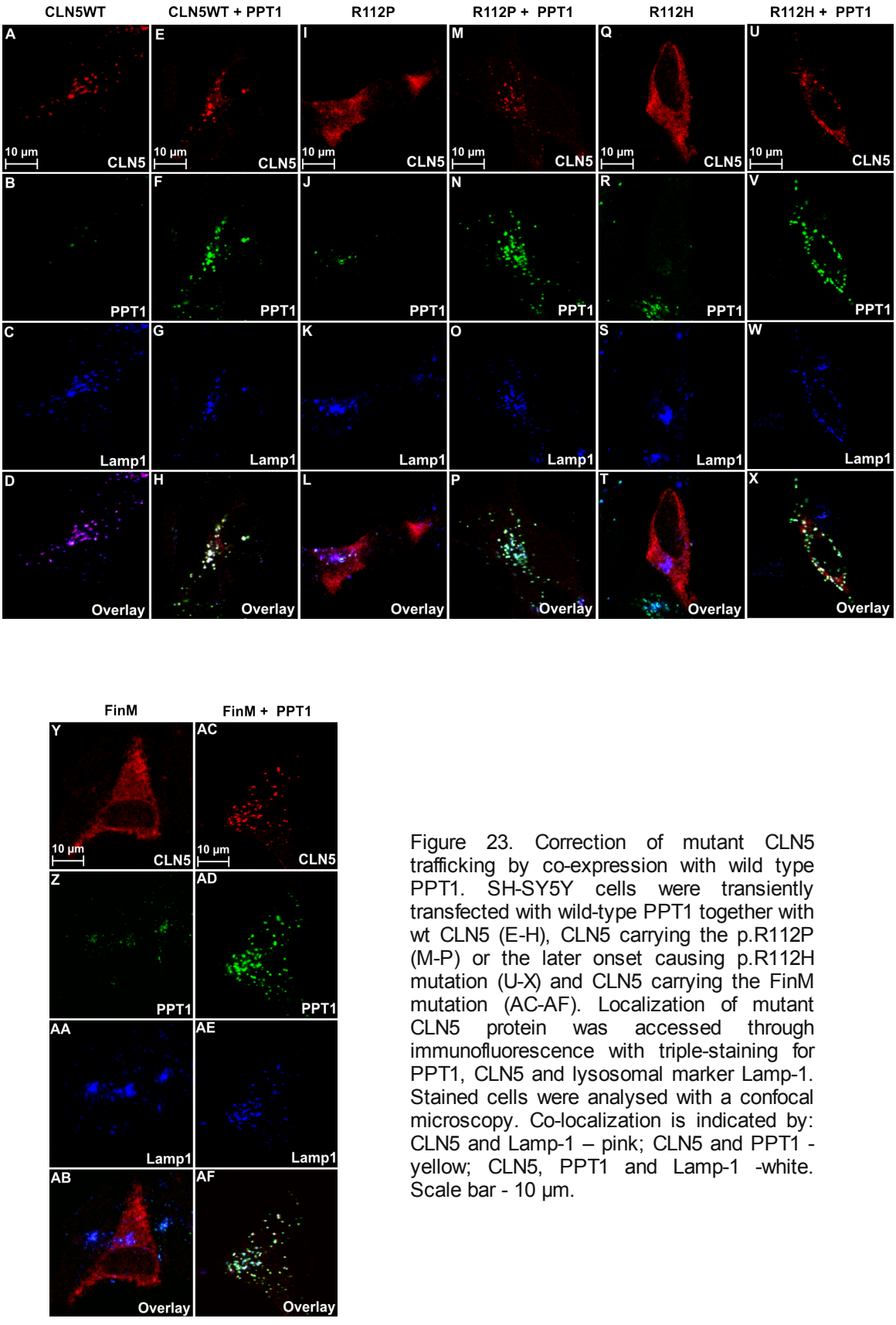


Figure 23. Correction of mutant CLN5 trafficking by co-expression with wild type PPT1. SH-SY5Y cells were transiently transfected with wild-type PPT1 together with wt CLN5 (E-H), CLN5 carrying the p.R112P (M-P) or the later onset causing p.R112H mutation (U-X) and CLN5 carrying the FinM mutation (AC-AF). Localization of mutant CLN5 protein was accessed through immunofluorescence with triple-staining for PPT1, CLN5 and lysosomal marker Lamp-1. Stained cells were analysed with a confocal microscopy. Co-localization is indicated by: CLN5 and Lamp-1 – pink; CLN5 and PPT1 - yellow; CLN5, PPT1 and Lamp-1 -white. Scale bar - 10 µm.

10 General discussion and conclusions

Neuronal Ceroid LIpofuscinoses (NCLs) are a group of heterogeneous neurological disorders. Clinically symptoms comprise loss of vision, epilepsy, progressive mental retardation and a reduced lifespan. These diseases are characterized by ubiquitous accumulation of an autofluorescent lipopigment in patient's tissues; however, the central nervous system is the most dramatically affected organ. Although certain forms of NCL are rare and regionally concentrated, as a whole NCLs are considered to be among the most common hereditary disease of childhood. These diseases result from mutations in distinct CLN genes (*CLN1-CLN10*) dispersed by all over the human genome. Eight of these genes have already been cloned (*CLN1-3*, *CLN5-8*, and *CLN10*) (Jalanko and Braulke, 2008).

A large portion of NCL patients identified worldwide presents either the classical or the variant subtype of the late-infantile phenotype. Likewise, in the Portuguese population LINCL and vLINCL were found to represent approximately 10 % and 42 % of all NCL patients identified thus far (Teixeira *et al*., 2003). As several genes are known to underlying these clinical phenotypes (*CLN2*, *CLN5*, *CLN6*, *CLN7* and *CLN10*) the knowledge on the genetic basis of these variants is not always straightforward. However, such knowledge is important for a better understanding of the molecular mechanisms behind the disease. Aiming to contribute to extend the information on this particular field, two diseases, CLN2 and CLN5, were studied by integrating data from DNA, RNA and protein analysis.

10.1 LINCL: evidence for mild disease due to an intron retention defect

The classical form of LINCL is usually observed in patients with mutations in the *CLN2* gene. This gene was identified in 1997 (Sleat *et al*., 1997) and since then numerous mutations leading to LINCL have been identified in patients with a deficient *in vitro* activity of TPP1 (Siintola *et al*., 2006a). TPP1 is a classical soluble lysosomal protease and although the specific *in vivo* substrate(s) are still unknown it is one of the most comprehensively studied NCL proteins. Worldwide, 66 disease causing mutations have been described in the *CLN2* gene, 12 of which are reported as splicing mutations (NCL

mutation database). Although *CLN2* mutations typically result in a late infantile phenotype (LINCL), rare splicing mutations in the *CLN2* gene have been described in association with milder clinical phenotypes, either vLINCL or JNCL (Sleat *et al*., 1999). Characterization of the molecular mechanisms of splice site mutations, including the ones occurring outside of the well conserved splice sites, is crucial for the establishment of genotype-phenotype correlations and, consequently, for the improvement of the genetic counseling. Moreover, implications for future therapeutic approaches may also arise as the existence of a low amount of a functionally active soluble protein may have therapeutic consequences distinct from those presented by the complete absence of enzyme.

In this study we have identified a novel splice site mutation, IVS7-10G>A (g.4196A>G), in the *CLN2* gene of a patient presenting an atypical later LINCL phenotype. Aiming to understand how the intronic mutation found in the Portuguese patient is compatible with the observed mild phenotype, several studies were conducted to characterize its exact molecular mechanism.

The novel mutation described here occurs in a stretch of adenines proximal to the wild-type 3' splice site in exon 7 of the *CLN2* gene. *In silico* analysis of the primary nucleotide sequence suggests that the mutation creates a new acceptor splice site in the intron 7 of the gene. Although the estimated strength of the newly created 3'-ss is higher than the wild-type 3'-ss, it is still overall weak and so, theoretically, both sites could be competing in the splicing process. In order to evaluate which transcripts are produced in the presence of the mutation several experimental approaches were pursued. Northern blot analysis and qRT-PCR of the RNA from the patient's fibroblasts showed no evidence of altered processing and steady-state level of the *CLN2*/mRNA. Sequencing of the cDNA showed the existence of one unique transcript, longer than the wild-type, resulting from the retention of the last 9 nucleotides of intron 7. Taken together, these observations suggest that only the new 3'-ss in intron 7 is used by the splicing machinery. Furthermore, the experimental data is in agreement with the *in silico* analysis, which in addition to assign a higher score to the novel 3'-ss also predicts the creation of new exonic splice enhancers in the presence of the mutation. Moreover, the usage of the novel 3'-ss complies with the scanning model for the selection of the 3' splice site (Smith *et al*., 1993), which proposes that the first "AG" downstream of the branch point and polypyrimidine tract is used by the spliceosome.

The IVS7-10G>A (g.4196A>G) mutation resulted in the production of a polypeptide with 3 additional amino acids which could impair the protein stability and/or maturation,

and, thus, affect the enzyme biological activity. The study of the endogenous CLN2 protein from patient's fibroblasts by Western blotting confirmed the presence of a polypeptide with an apparently correct maturation. However, a small reduction in the protein level was observed. Moreover, the *in vitro* enzymatic activity of the mutant TPP1 was found to be severely impaired, representing approximately 2-4 % of control activity in either leukocyte or fibroblast cells. In fact, the insertion of three amino acids (p.Pro295_Gly296insGluAsnPro) occurs between conserved amino acids proline295 and glycine296 located in relative proximity to two residues (Glu77 and Asp81) of the catalytic triad (Ser280-Glu77-Asp81) of the mature CLN2p (Wlodawer *et al*., 2003; Oyama *et al*., 2005). It is likely that the insertion of 3 aa at this position could lead to structural alterations that hinder protein activity, thus causing disease. However, although the enzymatic activity was significantly impaired, it was approximately 2-fold higher than that observed in classical LINCL patient cells carrying the p.R208X mutation. The p.R208X mutation results in the absence of TPP1 and has been previously associated with the classical, severe LINCL phenotype (Teixeira *et al*., 2003; NCL Mutation Database). Thus, the higher residual activity observed for the IVS7-10A>G patient is therefore likely to be responsible for the late onset and delayed progression of the disease observed in the Portuguese patient.

It is generally accepted that the amount of enzyme activity is inversely correlated to the phenotype severity. Moreover, it has been demonstrated that restoration of relatively low levels of residual activity can have a profound therapeutic benefit for most LSDs (Sands and Davidson, 2006). Our data reinforced this notion as it provided evidence that a residual activity of TPP1, albeit still very reduced, can have a great impact in the clinical spectra and longevity of the patient. In the specific case of TPP1, a recent study using mouse *CLN2* mutant models has reached a similar conclusion (Sleat *et al*., 2008). The authors demonstrated that slight increases in the residual activity of TPP1 can delay the disease onset and slow the progression of the phenotype. Altogether, these observations could provide valuable hindsight in the development of future therapies for the CLN2/TPP1 deficiency, specially for those based on the restoration of minimum enzyme activity sufficient to maintain a nearly normal cell homeostasis.

10.2 Insights into the molecular mechanisms of CLN5 deficiency

Mutations in the *CLN5* gene, identified and characterized in 1998 (Savukoski *et al*.), are the molecular basis of the Finnish variant LINCL. This LINCL subtype is extremely rare worldwide but particularly prevalent in Finland (Santavuori *et al*., 2000). At the time this study was initiated only five mutations had been identified in the *CLN5* gene, and only two described outside northern Europe (NCL Mutation Database).

A Portuguese NCL patient presenting a clinical phenotype suggestive of variant late-infantile NCL was screened for mutations in the *CLN5* gene. Analysis of the coding sequence and intron boundaries revealed the presence of two novel mutations (c.335G>C /p.R112P, c.565C>T/p.Q189X) in addition to the c.835G>A/p.D279N mutation. The p.D279N mutation had already been described before in a Dutch homozygous patient with a similar phenotype to the one observed in the Portuguese case (Savukoski *et al*., 1998), and is frequently referred to as the European (EUR) mutation. The patient was found to be a compound heterozygous, with the nonsense mutation in the paternal allele and both missense mutations in the allele inherited from the mother. Such complex alleles have been described in other diseases (Kappler *et al*., 1994; Shroyer *et al*., 2000). This represented the first evidence of the Finnish vLINCL in the Portuguese population and elevated the number of known mutations at the time to 7.

Both p.R112P and p.D279N mutations occur in evolutionary conserved residues and as such are likely to result in alterations that may impair the structure or the function of the protein. The novel p.R112P mutation occurs at the same residue as the p.R112H mutation, identified in homozygosity in Colombian patients (Pineda-Trujillo *et al*., 2005), suggesting an important structural or functional role for the R112 amino acid. However, while the c.335G>A mutation results in a change of an arginine for an also polar charged histidine amino acid (p.R112H), in the case of the c.335G>C mutation it involves the substitution of arginine with the non polar residue proline (p.R112P), which due to the cyclic structure of its side chain possesses more conformational constrains when compared to other amino acids. The novel c.565C>T (p.Q189X) mutation creates a stop codon at position 189 resulting most likely in a null allele.

Nonsense mutations often give rise to transcripts that are prematurely degraded by the cell through the quality control mechanism of nonsense-mediated mRNA decay (Frischmeyer and Dietz, 1999). To understand if this would be the case of the nonsense mutation identified in our study, CLN5/mRNA and the endogenous protein levels of patient's cultured fibroblasts were analysed. The steady state level of *CLN5*/mRNA in the

patients fibroblasts revealed a reduction to approximately 45 % of the control, confirming the prediction about the null nature of the c.565C>T allele. As such, the detected *CLN5*/mRNA must be originated from the complex allele, while the transcript resulting from the c.565C>T mutation is most likely degraded through nonsense-mediated mRNA decay. Consistently with this finding, a reduction of approximately 50% was also observed for the detected CLN5 polypeptides.

In the case of the Portuguese CLN5 patient, the establishment of genotype-phenotype correlations is not straightforward. This is mostly due to the presence of a complex allele carrying two mutations. Comparison with the p.R112H homozygous patients reported approximately at the same time in the Colombian population is also not direct. The Colombian patients described exhibit a juvenile onset of the disease, probably derived from the less deleterious amino acid alteration involved in the p.R112H mutation when compared with the p.R112P mutation. Moreover, the Portuguese patient carries an additional missense mutation, and a nonsense mutation in the other allele, which should account for the typical late infantile phenotype observed.

The notion that NCL proteins may be involved in a common unknown biochemical pathway has been suggested by several studies. Analysis of TPP1/CLN2 enzymatic activity for instance, has shown it to be elevated in some NCL subtypes such as CLN3 and CLN5 (Sleat, *et al*., 1998; Vesa *et al*., 2002). A recent and comprehensive study has also described interactions between several NCL proteins, including CLN1, CLN2, CLN3, CLN6 and CLN8 (Persaud-Sawin *et al*., 2007). In the case of the CLN5 protein, it has been reported to co-immunoprecipitate with the CLN2 and CLN3 proteins (Vesa *et al*., 2002). Interactions at the protein level raises the possibility of transcriptional regulation in the corresponding protein-encoding genes, and in the context of the newly identified Portuguese vLINCLFin patient we investigated the possibility of co-regulation of expression between these genes. Therefore, using real-time RT-PCR, the level of *CLN5*, *CLN3* and *CLN2* mRNA(s) were analysed in fibroblasts from patients deficient in these proteins. The results showed a significant reduction of *CLN3*/mRNA in CLN2 and CLN5 deficient fibroblasts, and a small diminution of the level of CLN5/mRNA in CLN3 deficient cells. However, reciprocal changes in expression were not observed, suggesting that CLN2 and CLN5 may act upstream of CLN3 in a common biochemical cascade. Altogether, these data strengthens the hypothesis that NCL proteins may be connected through a common metabolic pathway.

10.3 Novel insights into the biological properties of CLN5

The human *CLN5* gene encodes a 407 amino acid polypeptide with a predicted molecular weight of 46 kDa. CLN5 shares no homology to previously reported proteins and is moderately conserved in mammals (Savukoski *et al*., 1998). One of the most interesting aspects of the CLN5 protein is the possibility of alternative translation initiation of the human protein as it possesses an N-terminal extremity with four initiation methionines (M1-4). These AUG codons have been shown to be used *in vitro* and *in vivo*, resulting in the production of four 39-47 kDa polypeptides (Isosomppi *et al*., 2002; Vesa *et al*., 2002).

Initial *in silico* analysis using the human CLN5 sequence starting from the first methionine predicted the existence of two transmembrane domains (Savukoski *et al*., 1998). However, the CLN5 sequence starting from the fourth methionine (M4) shares a great homology to the mouse Cln5 polypeptide, which has been described as a soluble lysosomal protein (Holmberg *et al*., 2004). We have tried to address this apparent discrepancy by re-evaluating the *in silico* predictions for all M1-M4 sequences using different theoretical models. Comparison of the data obtained with two theoretical models, SignalP (Nielsen *et al*., 1997; Bendtsen *et al*., 2004) and SigCleave (von Heijne, 1986; 1987), suggests that all M1-M4 polypeptides should contain an ER signal peptide and the corresponding cleavage site. The negative result obtained with SignalP for the CLN5-M1 polypeptide is most likely due to the distance of the signal peptide region to the N-terminus of the protein, which interferes with the cut-off value used by this model. This hypothesis was tested through the overexpression of N-terminal GFP-tagged human M1-M4 CLN5 constructs in HeLa and COS-1 cells. Interestingly, in all analysed cell models, and regardless of the initiation codon present, only one CLN5 polypeptide of 60 kDa was observed Western blot analysis using the novel C-terminal CLN5 specific antibody (C/32), thus confirming previous predictions. Moreover, GFP and CLN5 were detected in distinct cell compartments, CLN5 was shown to localize to the endosomal/lysosomal system, while GFP presented an ER-like distribution. This observation was also consistent with a cleaved N-terminal GFP-tagged signal peptide. Thus, translation of human CLN5 *in vivo* is mostly likely started from the fourth methionine, generating a polypeptide with similar to the mouse Cln5 protein and containing only one starting methionine. Altogether, these data suggest that the CLN5 human protein does undergo signal peptide cleavage upon translocation into the ER. Moreover, in agreement with previous observations (Isosomppi *et al*., 2002; Vesa *et al*., 2002), in transient expression cell systems all polypeptides were detected in the endosomal/lysosomal compartment. Furthermore, the detection of a

soluble 60 kDa CLN5 polypeptide has also been described by Isosomppi and colleagues in overexpressing BHK-21 cells (Isosomppi *et al*., 2002). Of notice, in COS-1 cells transfected with the human *CLN5* cDNA presenting only the fourth initiation AUG codon, a soluble polypeptide was also detected (Holmberg *et al*., 2004).

Our observations, however, seem to contradict a report that describes the human CLN5 translated from the first initiation codon to be membrane bound (Vesa *et al*., 2002). Additionally, previous analysis of endogenous CLN5 polypeptides from HSF of the Portuguese vLINCL$_{Fin}$ patient with the ACP33 antibody also suggested the existence of integral membrane polypeptides (chapter 9.2.2).

Several hypotheses could account for these apparent inconsistencies. The membrane bound CLN5 polypeptide referred by Vesa and colleagues was detected with an N-terminal specific antibody which could, theoretically, recognize only the non-cleaved CLN5 polypeptide(s). With regard to the polypeptides detected by the ACP33 antibody, it is possible that they represent yet unidentified splice variants. In fact, Northern blot analysis of multiple human tissues has lead to the detection of four *CLN5* transcripts, speculated to correspond to alternative spliced transcripts (Savukoski *et al*., 1998). Another possibility is related to the second predicted hydrophobic region on the CLN5 polypeptide, located to the C-terminal portion of the protein (Savukoski *et al*., 1998). After the removal of the first hydrophobic domain by SPaseI cleavage, the second hydrophobic region, although weak, could facilitate the membrane association of the protein.

The maturation and processing of the CLN5 protein was further studied in the scope of Publication-III (Appendix) by co-author Mia-Lisa Schmiedt. The use of stable overexpressing CLN5-SH-SY5Y cells and the novel C/32 antibody allowed for the identification of two CLN5 polypeptides, of 50 and 60 kDa, both observed in the endosomal/lysosomal system. Further analysis of the polypeptides revealed the presence of high-mannose and complex type oligosaccharides, and this observation is in agreement with previous reports (Isosomppi *et al*., 2002). Additionally, after complete removal of oligosaccharides by PNGaseF digestion, two distinct polypeptides were still visible, raising the possibility for additional post-translational modifications. Phosphorylation is such a candidate since the CLN5 amino acid sequence includes several putative sites for this modification. The maturation of CLN5 polypeptides was additionally investigated to determine the origin of the 50 and 60 kDa polypeptides. Overall, the data obtained by several approaches suggest that the 50 kDa polypeptide derives from the 60 kDa form at a post-Golgi location, most likely the lysosomes.

The data presented in this thesis give support to the notion of CLN5 as a soluble

lysosomal glycoprotein, as it contains high mannose type oligosaccharides. In the past, targeting of CLN5 to the lysosomes by the M6P receptor pathway has also been explored by other authors using several experimental approaches (Vesa *et al*., 2002; Isosomppi *et al*., 2002; Sleat *et al*., 2005; 2006a; 2007; 2009). Of notice, Sleat and colleagues (Sleat *et al*., 2006b) have also identified three M6P tagged residues in the CLN5 protein, affinity-purified using immobilized Man6-P receptors, from human and mouse brain samples. The M6P receptor pathway is the most common lysosomal trafficking pathway used by soluble lysosomal enzymes. However, the evidence for the presence of complex type oligossacharides upon treatment with PNGaseF led us to investigate the existence of alternative cell routes (Appendix). To that end, overexpressed mouse Cln5 and human CLN5 were analysed in MPR mouse deficient fibroblasts. In this cell model, both proteins were found to co-localize with M6P independent lysosomal protein markers. This novel finding suggests that CLN5 can also be targeted to lysosomes by a M6P independent pathway. As other lysosomal proteins have been described to utilize alternative routes for lysosomal targetting, two molecules come up as possible candidates for their involvement in the M6P-independent lysosomal transport of CLN5: sortilin and the lysosomal integral membrane protein type 2 (LIMP-2). The multiligand receptor sortilin has been reported to mediate the lysosomal transport of prosaposin and acid sphingomyelinase (Lefrancois *et al*., 2003; Ni and Morales, 2006), as well as cathepsin H and NCL protein cathepsin D (Canuel *et al*., 2008). For instance, the lysosomal enzyme defective in Gaucher disease, beta-glucocerebrosidase, binds LIMP-2 and proceeds to traffic to the lysosomes in an M6P-receptor pathway independent route. Whatever mechanism underlies this novel CLN5 transport pathway, its characterization could be important to the identification of the biological function of the CLN5 protein. Moreover, understanding of trafficking mechanisms aid in the development of future therapies, as enzyme delivery is a major issue in therapies concerning brain affecting diseases.

10.4 Impact of mutation in the cell biology of CLN5

The study of disease causing mutations and their molecular consequences is useful for the characterization of functionally or structurally important amino acids or proteins domains. In addition it provides a better understanding of genotype-phenotype correlations and could help to predict future disease progression. In the specific case of the CLN5 protein, whose structure and function have not yet been documented, this issue assumes special relevance.

Mutations in the *CLN5* gene have been described to result in polypeptides which localize to the ER (p.L358A*fs*; p.W379C)(Lebrun *et al*., 2009) and Golgi apparatus (p.Y392X)(Isosomppi *et al*., 2002), but also in proteins with correct lysosomal trafficking (p.Y392X;p.D297N)(Vesa *et al*., 2002). We have analysed the pathological mechanisms of the most relevant mutations in the *CLN5* gene in the light of the new data obtained with the C/32 antibody. Overall, the studied mutations (p.R112H, p.R112P, p.E253X, p.D297N and p.Y392X) resulted in the ER retention of the mutant CLN5 polypeptides, with the exception of the p.D279N mutation, which gives rise to a mutant polypeptide compatible with a lysosomal localization in a small percentage of cells. The discrepancies with regard to the previously reported studies are most likely the result of the diversity of cell types used in the overexpression procedures. The ER localization of the mutant polypeptides suggest that misfolding could be the cause for the ER retention. Surprisingly, further analysis of p.R112P, p.R112H and p.D279N mutant polypeptides revealed sensitivity to EndoH and PGNaseF digestion. This observation suggests that these mutant polypeptides contain both high-mannose and complex type sugars, indicating previous localization at the Golgi complex. Furthermore, this may imply that a retrograde transport event from the Golgi to the ER may be necessary for correct CLN5 processing/maturation. Moreover, it could explain why previous studies have reported partial ER localization for CLN5 polypeptides (Holmberg *et al*., 2004; Isosomppi *et al*., 2002).

With regard to genotype-phenotype correlations, both p.R112P and p.D279N mutations are associated with the typical $vLINCL_{FIN}$ phenotype, while the mutation p.R112H was reported in patients with a milder form of the disease. As mutant polypeptides have been detected in the ER and showed identical migration pattern after EndoH and PGNaseF digestion, these mutations possibly hit amino acids important for the function of CLN5 or disturb relevant molecular interactions of the protein. Supporting this notion, localisation of the mutated polypeptides does not seem to correlate with the severity of the disease phenotype. Moreover, the fact that mutant polypeptide carrying the substitution p.R112H was also observed in the ER nevertheless the homozigosity for this

mutation be associated with a later onset form of the disease, may imply that CLN5 functions already in the early secretory pathway. Interestingly, NCL proteins CLN6 and CLN8 have been shown to localize to the endoplasmic reticulum (Heine *et al*., 2007; Lonka *et al*., 2004; Mole *et al*., 2004), raising the hypothesis of a possible interaction with CLN5.

Vesa and colleagues were the first to describe NCL protein interactions, between CLN5 and CLN2/CLN3, by the use of co-immunoprecipitation and *in vitro* binding assays (Vesa et al., 2002). More recently, a report has described that CLN5 interacts with several other NCL proteins: PPT1, TPP1, CLN3, CLN6 and CLN8 (Lyly *et al*., in press). Furthermore, the authors have shown that overexpression of PPT1, and not other NCL proteins, can facilitate the lysosomal trafficking of mutant p.Y392X CLN5 protein. In light of this knowledge, we have tried to evaluate if overexpression of PPT1 could also facilitate the trafficking of CLN5 polypeptides resulting from the missense mutations p.R112P and p.R112H. We found PPT1 was able to “rescue” both analysed mutant CLN5 polypeptides from retention in the ER and facilitate their trafficking to the lysosomal system. Of notice, lysosomal trafficking of PPT1 was not affected by the presence of abundant overexpressed mutant CLN5. The molecular basis for the interaction that seems to be occurring between CLN5 and CLN1 is unkown at the moment. One may speculate that PPT1 could be acting as a chaperone and/or facilitator of the CLN5 trafficking. The later is specially interesting as our data suggested that CLN5 can be targeted to the lysosomes via a M6P receptor independent pathway and CLN1 itself has been proposed to possess diferent properties from those of classical lysosomal enzymes transported via the M6P pathway (Lyly *et al*., 2007). Functional implications for the CLN5 protein can also arise. PPT1/CLN1 has been shown to interact with the F(1)-complex of the mitochondrial ATP synthase (Lyly *et al.*, 2008), known to function as a receptor for apolipoprotein A-1 (Fabre *et al.*, 2006; Martinez *et al*., 2003), and as a result CLN5 could also be directly/indirectly connected to the maintenance of lipid homeostasis. As a matter of fact, lipid disturbances are a common theme in several NCL subtypes (Rusyn et al., 2008), and some NCL proteins have been suggested to have direct interaction with lipids or act in lipid related processes. For instance CLN3 as been shown to bind galactosylceramide and is proposed to play a role in the ER/Golgi to membrane raft transport of GalCer (Rusyn *et al*., 2008), and another NCL protein, CathepsinD/CLN10, has been proposed to regulate ABCA1-mediated lipid efflux (Haidar *et al*., 2006).

Overall these data give strength to the hypothesis that NCL proteins may play a role in a common yet unknown biochemical pathway. Nevertheless, a more profound

characterization of NCL proteins and depiction of the nature of CLN-CLN interactions, and other possible interaction partners, represent key points to fully understand the entire NCL system.

11 References

Ahtiainen, L., J. Kolikova, A. Mutka, K. Luiro, M. Gentile, E. Ikonen, L. Khiroug, A. Jalanko, and O. Kopra. 2007. Palmitoyl protein thioesterase 1 (Ppt1)-deficient mouse neurons show alterations in cholesterol metabolism and calcium homeostasis prior to synaptic dysfunction. *Neurobiol Dis*, 28(1), 52-64.

Ahtiainen, L., O. Van Diggelen, A. Jalanko, and O. Kopra. 2003. Palmitoyl protein thioesterase 1 is targeted to the axons in neurons. *The Journal of Comparative Neurology,* 455(3), 368-77.

Awano, T., M. L. Katz, D. P. O'Brien, I. Sohar, P. Lobel, J. R. Coates, S. Khan, G. C. Johnson, U. Giger, and G. S. Johnson. 2006. A frame shift mutation in canine TPP1 (the ortholog of human CLN2) in a juvenile Dachshund with neuronal ceroid lipofuscinosis. *Molecular Genetics and Metabolism,* 89(3), 254-60.

Awano, T., M. L. Katz, D. P. O'Brien, J. F. Taylor, J. Evans, S. Khan, I. Sohar, P. Lobel, and G. S. Johnson. 2006. A mutation in the cathepsin D gene (CTSD) in American Bulldogs with neuronal ceroid lipofuscinosis. *Molecular Genetics and Metabolism,* 87(4), 341-8.

Bendtsen, J., Nielsen, H., von Heijne, G., Brunak, S., 2004. Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol*., 340, 783-795.

Berkovic, S. F., S. Carpenter, F. Andermann, E. Andermann, and L. S. Wolfe. 1988. Kufs' disease: a critical reappraisal. *Brain: A Journal of Neurology,* 111(Pt 1), 27-62.

Bernardini, F., and Warburton, M.J., 2002. Lysosomal degradation of cholecystokinin-(29-33)-amide in mouse brain is dependent on tripeptidyl peptidase-I: implications for the degradation and storage of peptides in classical late-infantile neuronal ceroid lipofuscinosis. *The Biochemical Journal*, 366(Pt 2), 521-9.

Bessa, C., C. A. F. Teixeira, M. Mangas, A. Dias, M. C. Sá Miranda, A. Guimarães, J. C. Ferreira, N. Canas, P. Cabral, and M. G. Ribeiro. 2006. Two novel CLN5 mutations in a Portuguese patient with vLINCL: insights into molecular mechanisms of CLN5 deficiency. *Molecular Genetics and Metabolism,* 89(3), 245-53.

Bielschowsky, M. 1913. Ueber spätinfantile amaurotische Idiotie mit Kleinhirn-symptomen. *Dtsch. Z. Nervenheilkd*, 50, 7-29.

Bildfell, R., C. Matwichuk, S. Mitchell, and P. Ward. 1995. Neuronal ceroid-lipofuscinosis in a cat. *Veterinary Pathology,* 32(5), 485-8.

Bleistein, J., Heidrich, H.G., and Debuch, H., 1980. The phospholipids of liver lysosomes from untreated rats. *Hoppe-Seyler's Zeitschrift Für Physiologische Chemie*, 361(4), 595-7.

Bonifacino, J. S., and L. M. Traub. 2003. Signals for sorting of transmembrane proteins to endosomes and lysosomes. *Annual Review of Biochemistry,* 72: 395-447.

Braulke, T., Bonifacino, J. S., 2009. Sorting of lysosomal proteins. *Biochim Biophys Acta*.

1793(4):605-14.

Bright, N. A., M. J. Gratian, and J. P. Luzio. 2005. Endocytic delivery to lysosomes mediated by concurrent fusion and kissing events in living cells. *Current Biology,* 15(4), 360-5.

Bright, N. A., B. J. Reaves, B. M. Mullock, and J. P. Luzio. 1997. Dense core lysosomes can fuse with late endosomes and are re-formed from the resultant hybrid organelles. *Journal of Cell Science,* 110(Pt 17), 2027-40.

Bronson, R. T., B. D. Lake, S. Cook, S. Taylor, and M. T. Davisson. 1993. Motor neuron degeneration of mice is a model of neuronal ceroid lipofuscinosis (Batten's disease). *Annals of Neurology,* 33(4), 381-5.

Brück, W., and H. H. Goebel. 1998. Microgial activation in neuronal ceroid-lipofuscinosis, 5, 276.

Camp, L.A., and Hofmann, S.L., 1993. Purification and properties of a palmitoyl-protein thioesterase that cleaves palmitate from H-Ras. *The Journal of Biological Chemistry*, 268(30), 22566-74.

Camp, L.A., Verkruyse, L.A., Afendis, S.J., Slaughter, C.A., and Hofmann, S.L., 1994. Molecular cloning and expression of palmitoyl-protein thioesterase. *The Journal of Biological Chemistry*, 269(37), 23212-9.

Canuel, M., A. Korkidakis, K. Konnyu, and C. R. Morales. 2008. Sortilin mediates the lysosomal targeting of cathepsins D and H. *Biochemical and Biophysical Research Communications,* 373(2), 292-7.

Cartegni, L., S. L. Chew, A. R. Krainer. 2002. Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat. Rev. Genet*., 3, 285-298.

Cartegni, L., J. Wang, Z. Zhu, M. Q. Zhang, and A. R. Krainer. 2003. ESEfinder: a web resource to identify exonic splicing enhancers. *Nucleic Acids Research,* 31(3), 3568-3571.

Castaneda, J., M. Lim, J. Cooper, and D. Pearce. 2008. Immune system irregularities in lysosomal storage disorders. *Acta Neuropathologica,* 115(2), 159-174.

Chang, M., J. D. Cooper, D. E. Sleat, S. H. Cheng, J. C. Dodge, M. A. Passini, P. Lobel, and B. L. Davidson. 2008. Intraventricular Enzyme Replacement Improves Disease Phenotypes in a Mouse Model of Late Infantile Neuronal Ceroid Lipofuscinosis. *Molecular Therapy: The Journal of the American Society of Gene Therapy*, 16(4), 649-56.

Cismondi, A., R. Kohan, A. Ghio, and A. M. O. Ramirez. 2008. Novel human pathological mutations. *Human Genetics,* 123(5), 537-555.

Codlin, S., R. L. Haines, and S. E. Mole. 2008. btn1 Affects Endocytosis, Polarization of Sterol-Rich Membrane Domains and Polarized Growth in Schizosaccharomyces pombe. *Traffic,* 9(6), 936-950.

Cooper, J. D. 2003. Progress towards understanding the neurobiology of Batten disease

or neuronal ceroid lipofuscinosis. *Current Opinion in Neurology,* 16(2), 121-8.

Cooper, J.D., Russell, C., and Mitchison, H.M., 2006. Progress towards understanding disease mechanisms in small vertebrate models of neuronal ceroid lipofuscinosis. *Biochimica Et Biophysica Acta*, 1762(10), 873-89.

de Duve, C., Pressman, B.C., Gianetto, R., Wattiaux, R., and Appelmans, F., 1955. Tissue fractionation studies. 6. Intracellular distribution patterns of enzymes in rat-liver tissue. *Biochemical Journal*, 60(4), 604–617.

de Voer, G., van der Bent, P., Rodrigues, A.J.G., van Ommen, G.B., Peters, D.J.M., and Taschner, P.E.M., 2005. Deletion of the Caenorhabditis elegans homologues of the CLN3 gene, involved in human juvenile neuronal ceroid lipofuscinosis, causes a mild progeric phenotype. *Journal of Inherited Metabolic Disease*, 28(6), 1065-80.

Dhar, S., R. L. Bitting, S. N. Rylova, P. J. Jansen, E. Lockhart, D. D. Koeberl, A. Amalfitano, and RM. Boustany. 2002. Flupirtine blocks apoptosis in batten patient lymphoblasts and in human postmitotic CLN3- and CLN2-deficient neurons. *Annals of Neurology,* 51(4), 448-66.

Dittmer, F., E. J. Ulbrich, A. Hafner, W. Schmahl, T. Meister, R. Pohlmann, and K. von Figura. 1999. Alternative mechanisms for trafficking of lysosomal enzymes in mannose 6-phosphate receptor-deficient mice are cell type-specific. *Journal of Cell Science,* 112(Pt 10), 1591-1597.

Doray, B., P. Ghosh, J. Griffith, H. J. Geuze, and S. Kornfeld. 2002. Cooperation of GGAs and AP-1 in packaging MPRs at the trans-Golgi network. *Science,* 297(5587), 1700-3.

Du, P. G., S. Kato, Y. H. Li, T. Maeda, T. Yamane, S. Yamamoto, M. Fujiwara, Y. Yamamoto, K. Nishi, and I. Ohkubo. 2001. Rat tripeptidyl peptidase I: molecular cloning, functional expression, tissue localization and enzymatic characterization. *Biological Chemistry,* 382(12), 1715-25.

Dyken, P. R. 1988. Reconsideration of the classification of the neuronal ceroid-lipofuscinoses. *American Journal of Medical Genetics*, Supplement 5, 69-84.

Eskelinen, EL., Y. Tanaka, and P. Saftig. 2003. At the acidic edge: emerging functions for lysosomal membrane proteins. *Trends in Cell Biology,* 13(3), 137-45.

Ezaki, J., M. Takeda-Ezaki, and E. Kominami. 2000. Tripeptidyl peptidase I, the late infantile neuronal ceroid lipofuscinosis gene product, initiates the lysosomal degradation of subunit c of ATP synthase. *Journal of Biochemistry,* 128(3), 509-16.

Ezaki, J., M. Takeda-Ezaki, K. Oda, and E. Kominami. 2000. Characterization of endopeptidase activity of tripeptidyl peptidase-I/CLN2 protein which is deficient in classical late infantile neuronal ceroid lipofuscinosis. *Biochemical and Biophysical Research Communications,* 268(3), 904-8.

Ezaki, J., I. Tanida, N. Kanehagi, and E. Kominami. 1999. A Lysosomal Proteinase, the Late Infantile Neuronal Ceroid Lipofuscinosis Gene (*CLN2*) Product, Is Essential for Degradation of a Hydrophobic Protein, the Subunit c of ATP Synthase. *Journal of Neurochemistry,* 72(6), 2573-2582.

Fabre, A. C., Vantourout, P., Champagne, E., Terce, F., Rolland, C., Perret, B., Collet, X., Barbaras, R., Martinez, L.O., 2006. Cell surface adenylate kinase activity regulates the F(1)-ATPase/P2Y (13)-mediated HDL endocytosis pathway on human hepatocytes. *Cell Mol. Life Sci.,* 63, 2829–2837.

Fairbrother, W. G., RF. Yeh, P. A. Sharp, and C. B. Burge. 2002. Predictive identification of exonic splicing enhancers in human genes. *Science,* 297(5583), 1007-1013.

Frischmeyer, P. A., Dietz, H. C., 1999. Nonsense-mediated mRNA decay in health and disease, *Hum. Mol. Genet*.,8(10), 1893–1900.

Frugier, T., Mitchell, N.L., Tammen, I., Houweling, P.J., Arthur, D.G., Kay, G.W., van Diggelen, O.P., Jolly, R.D., and Palmer, D.N., 2008. A new large animal model of CLN5 neuronal ceroid lipofuscinosis in Borderdale sheep is caused by a nucleotide substitution at a consensus splice site (c.571+1G>A) leading to excision of exon 3. *Neurobiology of Disease*, 29(2), 306-15.

Fukuda, M. 1991. Lysosomal membrane glycoproteins. Structure, biosynthesis, and intracellular trafficking. *The Journal of Biological Chemistry,* 266(32), 21327-30.

Futerman, A. H., and G. van Meer. 2004. The cell biology of lysosomal storage disorders. *Nature Reviews. Molecular Cell Biology,* 5(7), 554-65.

Gachet, Y., S. Codlin, J. S. Hyams, and S. E. Mole. 2005. btn1, the Schizosaccharomyces pombe homologue of the human Batten disease gene CLN3, regulates vacuole homeostasis. *J Cell Sci*, 118(23), 5525-5536.

Gao, Hanlin, Rose-Mary N Boustany, Janice A Espinola, Susan L Cotman, Lakshmi Srinidhi, Kristen Auger Antonellis, Tammy Gillis, et al. 2002. Mutations in a novel CLN6-encoded transmembrane protein cause variant neuronal ceroid lipofuscinosis in man and mouse. *American Journal of Human Genetics,* 70(2), 324-35.

Goebel, H. H., and H. Braak. 1989. Adult neuronal ceroid-lipofuscinosis. *Clinical Neuropathology,* 8(3), 109-19.

Golabek, A. A., N. Dolzhanskaya, M. Walus, K. E. Wisniewski, and E. Kida. 2008. Prosegment of tripeptidyl peptidase I is a potent, slow-binding inhibitor of its cognate enzyme. *The Journal of Biological Chemistry,* 283(24), 16497-504.

Golabek, A. A., and E. Kida. 2006. Tripeptidyl-peptidase I in health and disease. *Biological Chemistry,* 387(8), 1091-9.

Golabek, A. A., E. Kida, M. Walus, P. Wujek, P. Mehta, and K. E. Wisniewski. 2003. Biosynthesis, Glycosylation, and Enzymatic Processing in Vivo of Human Tripeptidyl-peptidase I. *J. Biol. Chem.,* 278(9), 7135-7145.

Golabek, A. A., P. Wujek, M. Walus, S. Bieler, C. Soto, K. E. Wisniewski, and E. Kida. 2004. Maturation of human tripeptidyl-peptidase I in vitro. *The Journal of Biological Chemistry,* 279(30), 31058-67.

Griffey, M., E. Bible, C. Vogler, B. Levy, P. Gupta, J. Cooper, and M. S. Sands. 2004. Adeno-associated virus 2-mediated gene therapy decreases autofluorescent storage material and increases brain mass in a murine model of infantile neuronal

ceroid lipofuscinosis. *Neurobiology of Disease,* 16(2), 360-9.

Griffiths, G. 1996. On vesicles and membrane compartments. *Protoplasma,* 195(1): 37-58.

Guhaniyogi, Jayita, Istvan Sohar, Kalyan Das, Ann M. Stock, and Peter Lobel. 2009. Crystal Structure and Autoactivation Pathway of the Precursor Form of Human Tripeptidyl-peptidase 1, the Enzyme Deficient in Late Infantile Ceroid Lipofuscinosis. *The Journal of Biological Chemistry,* 284(6), 3985-3997.

Guillas, I., Kirchman, P.A., Chuard, R., Pfefferli, M., Jiang, J.C., Jazwinski, S.M., and Conzelmann, A., 2001. C26-CoA-dependent ceramide synthesis of Saccharomyces cerevisiae is operated by Lag1p and Lac1p. *The EMBO Journal*, 20(11), 2655–2665.

Gupta, P., A. A. Soyombo, A. Atashband, K. E. Wisniewski, J. M. Shelton, J. A. Richardson, R. E. Hammer, and S. L. Hofmann. 2001. Disruption of PPT1 or PPT2 causes neuronal ceroid lipofuscinosis in knockout mice. *Proceedings of the National Academy of Sciences of the United States of America,* 98(24), 13566-13571.

Haidar, B., R. S. Kiss, L. Sarov-Blat, R. Brunet, C. Harder, R. McPherson, and Y. L. Marcel. 2006. Cathepsin D, a Lysosomal Protease, Regulates ABCA1-mediated Lipid Efflux. *J. Biol. Chem.,* 281(52), 39971-39981.

Haltia, M. 2003. The neuronal ceroid-lipofuscinoses. *Journal of Neuropathology and Experimental Neurology,* 62(1), 1-13.

Haltia, M., 2006. The neuronal ceroid-lipofuscinoses: from past to present. *Biochimica Et Biophysica Acta*, 1762(10), 850-6.

Hartikainen, J. M., W. Ju, K. E. Wisniewski, D. N. Moroziewicz, A. L. Kaczmarski, L. McLendon, D. Zhong, C. T. Suarez, W. T. Brown, and N. Zhong. 1999. Late infantile neuronal ceroid lipofuscinosis is due to splicing mutations in the CLN2 gene. *Molecular Genetics and Metabolism,* 67(2), 162-8.

Heine, C., A. Quitsch, S. Storch, Y. Martin, L. Lonka, AE. Lehesjoki, S. E. Mole, and T. Braulke. 2007. Topology and endoplasmic reticulum retention signals of the lysosomal storage disease-related membrane protein CLN6. *Molecular Membrane Biology*, 24(1), 74 - 87.

Heinonen, O., A. Kyttälä, E. Lehmus, T. Paunio, L. Peltonen, and A. Jalanko. 2000. Expression of palmitoyl protein thioesterase in neurons. *Molecular Genetics and Metabolism,* 69(2), 123-9.

Heinonen, O., T. Salonen, A. Jalanko, L. Peltonen, and A. Copp. 2000. CLN-1 and CLN-5, genes for infantile and variant late infantile neuronal ceroid lipofuscinoses, are expressed in the embryonic human brain. *The Journal of Comparative Neurology,* 426(3), 406-12.

Hellsten, E., J. Vesa, V. M. Olkkonen, A. Jalanko, and L. Peltonen. 1996. Human palmitoyl protein thioesterase: evidence for lysosomal targeting of the enzyme and disturbed cellular routing in infantile neuronal ceroid lipofuscinosis. *The EMBO Journal,* 15(19), 5240-5245.

Hermansson, M., R. Käkelä, M. Berghäll, AE. Lehesjoki, P. Somerharju, and U. Lahtinen. 2005. Mass spectrometric analysis reveals changes in phospholipid, neutral sphingolipid and sulfatide molecular species in progressive epilepsy with mental retardation, EPMR, brain: a case study. *Journal of Neurochemistry,* 95(3), 609-17.

Hickey, A. J., H. L. Chotkowski, N. Singh, J. G. Ault, C. A. Korey, M. E. MacDonald, and R. L. Glaser. 2006. Palmitoyl-protein thioesterase 1 deficiency in Drosophila melanogaster causes accumulation of abnormal storage material and reduced life span. *Genetics,* 172(4), 2379-90.

Hobert, J. A., and G. Dawson. 2006. Neuronal ceroid lipofuscinoses therapeutic strategies: past, present and future. *Biochimica Et Biophysica Acta,* 1762(10), 945-53.

Hogue, D. L., C. Nash, V. Ling, and T. C. Hobman. 2002. Lysosome-associated protein transmembrane 4 alpha (LAPTM4 alpha) requires two tandemly arranged tyrosine-based signals for sorting to lysosomes. *The Biochemical Journal,* 365(Pt 3), 721-30.

Holmberg, V., A. Jalanko, J. Isosomppi, AL. Fabritius, L. Peltonen, and O. Kopra. 2004. The mouse ortholog of the neuronal ceroid lipofuscinosis CLN5 gene encodes a soluble lysosomal glycoprotein expressed in the developing brain. *Neurobiology of Disease,* 16(1), 29-40.

Holmberg, V., L. Lauronen, T. Autti, P. Santavuori, M. Savukoski, P. Uvebrant, I. Hofman, L. Peltonen, and I. Järvelä. 2000. Phenotype-genotype correlation in eight patients with Finnish variant late infantile NCL (CLN5). *Neurology,* 55(4), 579-81.

Höning, S., J. Griffith, H. J. Geuze, and W. Hunziker. 1996. The tyrosine-based lysosomal targeting signal in lamp-1 mediates sorting into Golgi-derived clathrin-coated vesicles. *The EMBO Journal,* 15(19), 5230-9.

Houweling, Peter J, Julie A L Cavanagh, David N Palmer, Tony Frugier, Nadia L Mitchell, Peter A Windsor, Herman W Raadsma, and Imke Tammen. 2006. Neuronal ceroid lipofuscinosis in Devon cattle is caused by a single base duplication (c.662dupG) in the bovine CLN5 gene. *Biochimica Et Biophysica Acta,* 1762(10), 890-7.

Huang, K., and A. El-Husseini. 2005. Modulation of neuronal protein trafficking and function by palmitoylation. *Current Opinion in Neurobiology,* 15(5), 527-35.

Hunziker, W., and H. J. Geuze. 1996. Intracellular trafficking of lysosomal membrane proteins. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology,* 18(5), 379-89.

Isosomppi, J., J. Vesa, A. Jalanko, and L. Peltonen. 2002. Lysosomal localization of the neuronal ceroid lipofuscinosis CLN5 protein. *Human Molecular Genetics,* 11(8), 885-91.

Jabs, S., A. Quitsch, R. Käkelä, B. Koch, J. Tyynelä, H. Brade, M. Glatzel, et al. 2008. Accumulation of bis(monoacylglycero)phosphate and gangliosides in mouse models of neuronal ceroid lipofuscinosis. *Journal of Neurochemistry,* 106(3), 1415-25.

Jalanko, A., and T. Braulke. 2008. Neuronal ceroid lipofuscinoses. *Biochimica Et*

Biophysica Acta, 1793(4), 697-709.

Jalanko, A., Vesa, J., Manninen, T., von Schantz, C., Minye, H., Fabritius, A., Salonen, T., Rapola, J., Gentile, M., Kopra, U., and Peltonen, L., 2005. Mice with Ppt1Deltaex4 mutation replicate the INCL phenotype and show an inflammation-associated loss of interneurons. *Neurobiology of Disease*, 18(1), 226-41.

Janský, J. 1908. Sur un cas jusqu'a présente non décrit de L'idiotie amaurotique familiále compliqueé une hypoplasie du cervelet. *Sborna lék*, 13, 165-196.

Janvier, K., and J. S. Bonifacino. 2005. Role of the endocytic machinery in the sorting of lysosome-associated membrane proteins. *Molecular Biology of the Cell,* 16(9), 4231-42.

Järplid, B., and Haltia, M., 1993. An animal model of the infantile type of neuronal ceroid-lipofuscinosis. *Journal of Inherited Metabolic Disease*, 16(2), 274-7.

Järvelä, I., Lehtovirta, M., Tikkanen, R., Kyttala, A., and Jalanko, A., 1999. Defective intracellular transport of CLN3 is the molecular basis of Batten disease (JNCL). *Hum. Mol. Genet.*, 8(6), 1091-1098.

Järvelä, I., Sainio, M., Rantamaki, T., Olkkonen, V.M., Carpen, O., Peltonen, L., and Jalanko, A., 1998. Biosynthesis and intracellular targeting of the CLN3 protein defective in Batten disease. *Hum. Mol. Genet.*, 7(1), 85-90.

Journet, A., Chapel, A., Kieffer, S., Roux, F., and Garin, J., 2002. Proteomic analysis of human lysosomes: application to monocytic and breast cancer cells. *Proteomics*, 2(8), 1026-40.

Junaid, M.A., Wu, G., and Pullarkat, R.K., 2000. Purification and Characterization of Bovine Brain Lysosomal Pepstatin-Insensitive Proteinase, the Gene Product Deficient in the Human Late-Infantile Neuronal Ceroid Lipofuscinosis. *Journal of Neurochemistry*, 74(1), 287-294.

Kabeya, Y., Mizushima, N., Ueno, T., Yamamoto, A., Kirisako, T., Noda, T., Kominami, E., Ohsumi, Y., and Yoshimori, T., 2000. LC3, a mammalian homologue of yeast Apg8p, is localized in autophagosome membranes after processing. *The EMBO Journal*, 19(21), 5720–5728.

Käkelä, R., Somerharju, P., and Tyynelä, J., 2003. Analysis of phospholipid molecular species in brains from patients with infantile and juvenile neuronal-ceroid lipofuscinosis using liquid chromatography-electrospray ionization mass spectrometry. *Journal of Neurochemistry*, 84(5), 1051-65.

Kalatzis, V., Cherqui, S., Antignac, C., and Gasnier, B., 2001. Cystinosin, the protein defective in cystinosis, is a H(+)-driven lysosomal cystine transporter. *The EMBO Journal*, 20(21), 5940-9.

Kappler, J., Sommerlade, H. J., von Figura, K., Gieselmann, V., 1994. Complex arylsulfatase A alleles causing metachromatic leukodystrophy, *Hum. Mutat.,* 4(2), 119–127.

Katz, M.L., Gao, C.L., Prabhakaram, M., Shibuya, H., Liu, P.C., and Johnson, G.S., 1997. Immunochemical localization of the Batten disease (CLN3) protein in retina.

Investigative Ophthalmology & Visual Science, 38(11), 2375-86.

Katz, M.L., Khan, S., Awano, T., Shahid, S.A., Siakotos, A.N., and Johnson, G.S., 2005. A mutation in the CLN8 gene in English Setter dogs with neuronal ceroid-lipofuscinosis. *Biochemical and Biophysical Research Communications*, 327(2), 541-7.

Kida, E., Wisniewski, K.E., and Connell, F., 1995. Topographic variabilities of immunoreactivity to subunit c of mitochondrial ATP synthase and lectin binding in late infantile neuronal ceroid-lipofuscinosis. *American Journal of Medical Genetics*, 57(2), 182-6.

Kim, Y., Ramirez-Montealegre, D., and Pearce, D.A., 2003. A role in vacuolar arginine transport for yeast Btn1p and for human CLN3, the protein defective in Batten disease. *Proceedings of the National Academy of Sciences of the United States of America*, 100(26), 15458–15462.

Koike, M., Shibata, M., Ohsawa, Y., Nakanishi, H., Koga, T., Kametaka, S., Waguri, S., Momoi, T., Kominami, E., Peters, C., Figura, K.V., Saftig, P., and Uchiyama, Y., 2003. Involvement of two different cell death pathways in retinal atrophy of cathepsin D-deficient mice. *Molecular and Cellular Neurosciences*, 22(2), 146-61.

Koike, M., Shibata, M., Waguri, S., Yoshimura, K., Tanida, I., Kominami, E., Gotow, T., Peters, C., Figura, K.V., Mizushima, N., Saftig, P., and Uchiyama, Y., 2005. Participation of Autophagy in Storage of Lysosomes in Neurons from Mouse Models of Neuronal Ceroid-Lipofuscinoses (Batten Disease). *The American Journal of Pathology*, 167(6), 1713–1728.

Kolter, T., and Sandhoff, K., 2006. Sphingolipid metabolism diseases. *Biochimica Et Biophysica Acta*, 1758(12), 2057-79.

Kominami, A.E., 2002. What are the requirements for lysosomal degradation of subunit c of mitochondrial ATPase? *IUBMB Life*, 54(2), 89-90.

Kopan, S., Sivasubramaniam, U., and Warburton, M.J., 2004. The lysosomal degradation of neuromedin B is dependent on tripeptidyl peptidase-I: evidence for the impairment of neuropeptide degradation in late-infantile neuronal ceroid lipofuscinosis. *Biochemical and Biophysical Research Communications*, 319(1), 58-65.

Koppang, N., 1988. The English setter with ceroid-lipofuscinosis: a suitable model for the juvenile type of ceroid-lipofuscinosis in humans. *American Journal of Medical Genetics. Supplement*, 5, 117-25.

Kopra, O., Vesa, J., von Schantz, C., Manninen, T., Minye, H., Fabritius, A., Rapola, J., van Diggelen, O., Saarela, J., Jalanko, A., and Peltonen, L., 2004a. A mouse model for Finnish variant late infantile neuronal ceroid lipofuscinosis, CLN5, reveals neuropathology associated with early aging. *Human Molecular Genetics*, 13(23), 2893-906.

Kopra, O., Vesa, J., von Schantz, C., Manninen, T., Minye, H., Fabritius, A., Rapola, J., van Diggelen, O.P., Saarela, J., Jalanko, A., and Peltonen, L., 2004b. A mouse model for Finnish variant late infantile neuronal ceroid lipofuscinosis, CLN5, reveals neuropathology associated with early aging. *Human Molecular Genetics*,

13(23), 2893-906.

Kornfeld, S., and William, S., 2001. I-Cell Disease and Pseudo-Hurler Polydystrophy: Disorders of Lysosomal Enzyme Phosphorylation and Localization. In *The Metabolic and Molecular Bases of Inherited Disease*. McGraw-Hill Professional, pp. 3469-3482.

Kremmidiotis, G., Lensink, I.L., Bilton, R.L., Woollatt, E., Chataway, T.K., Sutherland, G.R., and Callen, D.F., 1999. The Batten disease gene product (CLN3p) is a Golgi integral membrane protein. *Hum. Mol. Genet.*, 8(3), 523-531.

Kurachi, Y., Oka, A., Itoh, M., Mizuguchi, M., Hayashi, M., and Takashima, S., 2001. Distribution and development of CLN2 protein, the late-infantile neuronal ceroid lipofuscinosis gene product. *Acta Neuropathologica*, 102(1), 20-6.

Kyttälä, A., Lahtinen, U., Braulke, T., and Hofmann, S.L., 2006. Functional biology of the neuronal ceroid lipofuscinoses (NCL) proteins. *Biochimica Et Biophysica Acta*, 1762(10), 920-33.

Laemmli, U. K., Cleavage of structural proteins during the Assembly of the Head of Bacteriophage T4. 1970, *Nature*, 227, 680–685.

Lefrancois, S., Zeng, J., Hassan, A.J., Canuel, M., and Morales, C.R., 2003. The lysosomal trafficking of sphingolipid activator proteins (SAPs) is mediated by sortilin. *The EMBO Journal*, 22(24), 6430-7.

Lehtovirta, M., Kyttala, A., Eskelinen, E., Hess, M., Heinonen, O., and Jalanko, A., 2001. Palmitoyl protein thioesterase (PPT) localizes into synaptosomes and synaptic vesicles in neurons: implications for infantile neuronal ceroid lipofuscinosis (INCL). *Hum. Mol. Genet.*, 10(1), 69-75.

Lerner, T.J., Boustany, R., Anderson, J.W., D'Arigo, K.L., Schlumpf, K., Buckler, A.J., Gusella, J.F., and Haines, J.L., 1995. Isolation of a novel gene underlying batten disease, CLN3. *Cell*, 82(6), 949-957.

Lim, M.J., Beake, J., Bible, E., Curran, T.M., Ramirez-Montealegre, D., Pearce, D.A., and Cooper, J.D., 2006. Distinct patterns of serum immunoreactivity as evidence for multiple brain-directed autoantibodies in juvenile neuronal ceroid lipofuscinosis. *Neuropathology and Applied Neurobiology*, 32(5), 469-82.

Lin, L., Sohar, I., Lackland, H., and Lobel, P., 2001. The human CLN2 protein/tripeptidyl-peptidase I is a serine protease that autoactivates at acidic pH. *The Journal of Biological Chemistry*, 276(3), 2249-55.

Liu, C.G., Sleat, D.E., Donnelly, R.J., and Lobel, P., 1998. Structural organization and sequence of CLN2, the defective gene in classical late infantile neuronal ceroid lipofuscinosis. *Genomics*, 50(2), 206-12.

Lloyd, J.B., and Forster, S., 1986. The lysosome membrane. *Trends in Biochemical Sciences*, 11(9), 365-368.

Lonka, L., Kyttala, A., Ranta, S., Jalanko, A., and Lehesjoki, A., 2000. The neuronal ceroid lipofuscinosis CLN8 membrane protein is a resident of the endoplasmic reticulum. *Hum. Mol. Genet.*, 9(11), 1691-1697.

Lonka, L., Salonen, T., Siintola, E., Kopra, O., Lehesjoki, A., and Jalanko, A., 2004. Localization of wild-type and mutant neuronal ceroid lipofuscinosis CLN8 proteins in non-neuronal and neuronal cells. *Journal of Neuroscience Research*, 76(6), 862-71.

Lowry, O. H., Rosebrough, N.J., Farr, A.L., Randall, R.J., 1951. Protein measurement with the Folin phenol reagent, *J. Biol. Chem.,* 193, 265–275.

Luiro, K., Kopra, O., Lehtovirta, M., and Jalanko, A., 2001. CLN3 protein is targeted to neuronal synapses but excluded from synaptic vesicles: new clues to Batten disease. *Hum. Mol. Genet.*, 10(19), 2123-2131.

Luzio, J.P., Pryor, P.R., and Bright, N.A., 2007. Lysosomes: fusion and function. *Nature Reviews. Molecular Cell Biology*, 8(8), 622-32.

Lyly, A., von Schantz, C., Salonen, T., Kopra, O., Saarela, J., Jauhiainen, M., Kyttälä, A., Jalanko, A., 2007. Glycosylation, transport, and complex formation of palmitoyl protein thioesterase 1 (PPT1)--distinct characteristics in neurons. *BMC Cell Biol.*, 12, 8-22.

Lyly, A., von Schantz, C., Heine, C., Schmiedt, M-L., Sipilä, T., Jalanko A., Kyttälä, Aija., i*n press*. Novel interactions of CLN5 support molecular networking between Neuronal Ceroid Lipofuscinosis proteins. BMC Cell Biol.

Lyly, A., Marjavaara, S.K., Kyttälä, A., Uusi-Rauva, K., Luiro, K., Kopra, O., Martinez, L.O., Tanhuanpää, K., Kalkkinen, N., Suomalainen, A., Jauhiainen, M., and Jalanko, A., 2008. Deficiency of the INCL protein Ppt1 results in changes in ectopic F1-ATP synthase and altered cholesterol metabolism. *Human Molecular Genetics*, 17(10), 1406-17.

Marsh, M., Schmid, S., Kern, H., Harms, E., Male, P., Mellman, I., and Helenius, A., 1987. Rapid analytical and preparative isolation of functional endosomes by free flow electrophoresis. *The Journal of Cell Biology*, 104(4), 875-86.

Martinez, L. O., Jacquet, S., Esteve, J. P., Rolland, C., Cabezon, E., Champagne, E., Pineau, T., Georgeaud, V., Walker, J. E., Terce, F., et al. 2003. Ectopic beta-chain of ATP synthase is an apolipoprotein A-I receptor in hepatic HDL endocytosis. *Nature,* 421, 75–79

Matteoni, R., and Kreis, T.E., 1987. Translocation and clustering of endosomes and lysosomes depends on microtubules. *The Journal of Cell Biology*, 105(3), 1253-65.

van Meel, E., and Klumperman, J., 2008. Imaging and imagination: understanding the endo-lysosomal system. *Histochemistry and Cell Biology*, 129(3), 253-66.

Mellman, I., Fuchs, R., and Helenius, A., 1986. Acidification of the endocytic and exocytic pathways. *Annual Review of Biochemistry*, 55, 663-700.

Melville, S.A., Wilson, C.L., Chiang, C.S., Studdert, V.P., Lingaas, F., and Wilton, A.N.,

2005. A mutation in canine CLN5 causes neuronal ceroid lipofuscinosis in Border collie dogs. *Genomics*, 86(3), 287-94.

Michalak, M., Baksh, S., and Opas, M., 1991. Identification and immunolocalization of calreticulin in pancreatic cells: no evidence for "calciosomes". *Experimental Cell Research*, 197(1), 91-99.

Miller, S. A., Dykes, D. D., Polesky, H. F., 1988. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res*, 16(3), 1215.

Mitchison, H.M., Lim, M.J., and Cooper, J.D., 2004. Selectivity and types of cell death in the neuronal ceroid lipofuscinoses. *Brain Pathology*, 14(1), 86-96.

Mizutani, Y., Kihara, A., and Igarashi, Y., 2005. Mammalian Lass6 and its related family members regulate synthesis of specific ceramides. *Biochemical Journal*, 390(Pt1), 263–271.

Mole, S.E., Michaux, G., Codlin, S., Wheeler, R.B., Sharp, J.D., and Cutler, D.F., 2004. CLN6, which is associated with a lysosomal storage disease, is an endoplasmic reticulum protein. *Experimental Cell Research,* 298(2), 399-406.

Mullins, C., and Bonifacino, J.S., 2001. The molecular machinery for lysosome biogenesis. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, 23(4), 333-43.

Mullock, B.M., Bright, N.A., Fearon, C.W., Gray, S.R., and Luzio, J.P., 1998. Fusion of lysosomes with late endosomes produces a hybrid organelle of intermediate density and is NSF dependent. *The Journal of Cell Biology*, 140(3), 591-601.

Myllykangas, L., Tyynelä, J., Page-McCaw, A., Rubin, G.M., Haltia, M.J., and Feany, M.B., 2005. Cathepsin D-deficient Drosophila recapitulate the key features of neuronal ceroid lipofuscinoses. *Neurobiology of Disease*, 19(1-2), 194-9.

Narayan, S.B., Rakheja, D., Tan, L., Pastor, J.V., and Bennett, M.J., 2006. CLN3P, the Batten's disease protein, is a novel palmitoyl-protein Delta-9 desaturase. *Annals of Neurology*, 60(5), 570-7.

Nielsen, H., Engelbrecht, J., Brunak, S., von Heijne, G., 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Engineering, 10, 1-6.

Nijssen, P.C.G., Ceuterick, C., van Diggelen, O.P., Elleder, M., Martin, J., Teepen, J.L.J.M., Tyynelä, J., and Roos, R.A.C., 2003. Autosomal dominant adult neuronal ceroid lipofuscinosis: a novel form of NCL with granular osmiophilic deposits without palmitoyl protein thioesterase 1 deficiency. *Brain Pathology*, 13(4), 574-81.

Ni, X., and Morales, C.R., 2006. The Lysosomal Trafficking of Acid Sphingomyelinase is Mediated by Sortilin and Mannose 6-phosphate Receptor. *Traffic*, 7(7), 889-902.

Okouchi, M., Ekshyyan, O., Maracine, M., and Aw, T.Y., 2007. Neuronal apoptosis in neurodegeneration. *Antioxidants & Redox Signaling*, 9(8), 1059-96.

Oyama, H., Fujisawa, T., Suzuki, T., Dunn, B.M., Wlodawer, A., and Oda, K., 2005. Catalytic Residues and Substrate Specificity of Recombinant Human Tripeptidyl

Peptidase I (CLN2). *J Biochem*, 138(2), 127-134.

Pal, A., Kraetzner, R., Gruene, T., Grapp, M., Schreiber, K., Grønborg, M., Urlaub, H., Becker, S., Asif, A.R., Gärtner, J., Sheldrick, G.M., and Steinfeld, R., 2009. Structure of tripeptidyl-peptidase I provides insight into the molecular basis of late infantile neuronal ceroid lipofuscinosis. *The Journal of Biological Chemistry*, 284(6), 3976-3984.

Palmer, D.N., Fearnley, I.M., Walker, J.E., Hall, N.A., Lake, B.D., Wolfe, L.S., Haltia, M., Martinus, R.D., and Jolly, R.D., 1992. Mitochondrial ATP synthase subunit c storage in the ceroid-lipofuscinoses (Batten disease). *American Journal of Medical Genetics*, 42(4), 561-7.

Palmer, D.N., Husbands, D.R., Winter, P.J., Blunt, J.W., and Jolly, R.D., 1986. Ceroid lipofuscinosis in sheep. I. Bis(monoacylglycero)phosphate, dolichol, ubiquinone, phospholipids, fatty acids, and fluorescence in liver lipopigment lipids. *The Journal of Biological Chemistry*, 261(4), 1766-72.

Palmer, D.N., Martinus, R.D., Cooper, S.M., Midwinter, G.G., Reid, J.C., and Jolly, R.D., 1989. Ovine ceroid lipofuscinosis. The major lipopigment protein and the lipid-binding subunit of mitochondrial ATP synthase have the same NH2-terminal sequence. *The Journal of Biological Chemistry*, 264(10), 5736-40.

Pao, S.S., Paulsen, I.T., and Jr, M.H.S., 1998. Major Facilitator Superfamily. *Microbiology and Molecular Biology Reviews*, 62(1), 1–34.

Passini, M.A., Dodge, J.C., Bu, J., Yang, W., Zhao, Q., Sondhi, D., Hackett, N.R., Kaminsky, S.M., Mao, Q., Shihabuddin, L.S., Cheng, S.H., Sleat, D.E., Stewart, G.R., Davidson, B.L., Lobel, P., and Crystal, R.G., 2006. Intracranial delivery of CLN2 reduces brain pathology in a mouse model of classical late infantile neuronal ceroid lipofuscinosis. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 26(5), 1334-42.

Pearce, D.A., 2001. Experimental models of NCL: the yeast model. *Advances in Genetics*, 45, 205-16.

Pearce, D.A., and Sherman, F., 1998. A yeast model for the study of Batten disease. *Proceedings of the National Academy of Sciences of the United States of America*, 95(12), 6915-8.

Persaud-Sawin, D.A., and Boustany, R.N., 2005. Cell death pathways in juvenile Batten disease. *Apoptosis: An International Journal on Programmed Cell Death*, 10(5), 973-85.

Persaud-Sawin, D., Mousallem, T., Wang, C., Zucker, A., Kominami, E., and Boustany, R.N., 2007. Neuronal ceroid lipofuscinosis: a common pathway?, *Pediatr Res*, 61(2), 146-52.

Persaud-Sawin, D., VanDongen, A., and Boustany, R.N., 2002. Motifs within the CLN3 protein: modulation of cell growth rates and apoptosis. *Hum. Mol. Genet.*, 11(18), 2129-2142.

Peters, C., Braun, M., Weber, B., Wendland, M., Schmidt, B., Pohlmann, R., Waheed, A.,

and von Figura, K., 1990. Targeting of a lysosomal membrane protein: a tyrosine-containing endocytosis signal in the cytoplasmic tail of lysosomal acid phosphatase is necessary and sufficient for targeting to lysosomes. *The EMBO Journal*, 9(11), 3497-506.

Pfaffl, M.W., A new mathematical model for relative quantification in real-time PCR. 2001, *Nucleic Acids Res*., 29, 2002–2207.

Pierret, C., Morrison, J.A., and Kirk, M.D., 2008. Treatment of lysosomal storage disorders: focus on the neuronal ceroid-lipofuscinoses. *Acta Neurobiologiae Experimentalis*, 68(3), 429-42.

Pineda-Trujillo, N., Cornejo, W., Carrizosa, J., Wheeler, R.B., Munera, S., Valencia, A., Agudelo-Arango, J., Cogollo, A., Anderson, G., Bedoya, G., Mole, S.E., and Ruiz-Linares, A., 2005. A CLN5 mutation causing an atypical neuronal ceroid lipofuscinosis of juvenile onset. *Neurology*, 64(4), 740-742.

Pohlmann R., Boeker M.W., von Figura K. 1995. The two mannose 6-phosphate receptors transport distinct complements of lysosomal proteins. *J Biol Chem*, 270(45), 27311-8.

Press, E.M., Porter, R.R., and Cebra, J., 1960. The isolation and properties of a proteolytic enzyme, cathepsin D, from bovine spleen. *The Biochemical Journal*, 74, 501-14.

Puranam, K.L., Guo, W.X., Qian, W.H., Nikbakht, K., and Boustany, R.N., 1999. CLN3 defines a novel antiapoptotic pathway operative in neurodegeneration and mediated by ceramide. *Molecular Genetics and Metabolism*, 66(4), 294-308.

Rajawat, Y.S., and Bossis, I., 2008. Autophagy in aging and in neurodegenerative disorders. *Hormones*, 7(1), 46-61.

Rakheja, D., Narayan, S.B., Pastor, J.V., and Bennett, M.J., 2004. CLN3P, the Batten disease protein, localizes to membrane lipid rafts (detergent-resistant membranes). *Biochemical and Biophysical Research Communications*, 317(4), 988-91.

Ranta, S., Zhang, Y., Ross, B., Lonka, L., Takkunen, E., Messer, A., Sharp, J., Wheeler, R., Kusumi, K., Mole, S., Liu, W., Soares, M.B., Bonaldo, M.F., Hirvasniemi, A., de la Chapelle, A., Gilliam, T.C., and Lehesjoki, A.E., 1999. The neuronal ceroid lipofuscinoses in human EPMR and mnd mutant mice are associated with mutations in CLN8. *Nature Genetics*, 23(2), 233-6.

Rawlings, N.D., and Barrett, A.J., 1995. Families of aspartic peptidases, and those of unknown catalytic mechanism. *Methods in Enzymology*, 248, 105-20.

Reczek, D., Schwake, M., Schröder, J., Hughes, H., Blanz, J., Jin, X., Brondyk, W., Van Patten, S., Edmunds, T., and Saftig, P., 2007. LIMP-2 is a receptor for lysosomal mannose-6-phosphate-independent targeting of beta-glucocerebrosidase. *Cell*, 131(4), 770-83.

Riebeling, C., Allegood, J.C., Wang, E., Merrill, A.H., and Futerman, A.H., 2003. Two Mammalian Longevity Assurance Gene (LAG1) Family Members, trh1 and trh4, Regulate Dihydroceramide Synthesis Using Different Fatty Acyl-CoA Donors. *J.*

Biol. Chem., 278(44), 43452-43459.

Rouillé, Y., Rohn, W., and Hoflack, B., 2000. Targeting of lysosomal proteins. *Seminars in Cell & Developmental Biology*, 11(3), 165-71.

Rusyn, E., Mousallem, T., Persaud-Sawin, D., Miller, S., and Boustany, R.N., 2008. CLN3p impacts galactosylceramide transport, raft morphology, and lipid content. *Pediatric Research*, 63(6), 625-31.

Saftig, P., Hetman, M., Schmahl, W., Weber, K., Heine, L., Mossmann, H., Köster, A., Hess, B., Evers, M., and Figura, K.V., 1995. Mice deficient for the lysosomal proteinase cathepsin D exhibit progressive atrophy of the intestinal mucosa and profound destruction of lymphoid cells. *The EMBO Journal*, 14(15), 3599–3608.

Sands, M.S., and Davidson, B.L., 2006. Gene therapy for lysosomal storage diseases. *Molecular Therapy: The Journal of the American Society of Gene Therapy*, 13(5), 839-849.

Santavuori, P., Lauronen, L., Kirveskari, E., Aberg, L., Sainio, K., and Autti, T., 2000. Neuronal ceroid lipofuscinoses in childhood. *Neurological Sciences: Official Journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology*, 21(3 Suppl), S35-41.

Santavuori, P., Rapola, J., Nuutila, A., Raininko, R., Lappi, M., Launes, J., Herva, R., and Sainio, K., 1991. The spectrum of Jansky-Bielschowsky disease. *Neuropediatrics*, 22(2), 92-6.

Santavuori, P., Rapola, J., Sainio, K., and Raitta, C., 1982. A variant of Jansky-Bielschowsky disease. *Neuropediatrics*, 13(3), 135-41.

Savukoski, M., Klockars, T., Holmberg, V., Santavuori, P., Lander, E.S., and Peltonen, L., 1998. CLN5, a novel gene encoding a putative transmembrane protein mutated in Finnish variant late infantile neuronal ceroid lipofuscinosis. *Nature Genetics*, 19(3), 286-8.

Schantz, C.V., Saharinen, J., Kopra, O., Cooper, J.D., Gentile, M., Hovatta, I., Peltonen, L., and Jalanko, A., 2008. Brain gene expression profiles of Cln1 and Cln5 deficient mice unravels common molecular pathways underlying neuronal degeneration in NCL diseases. *BMC Genomics*, 9, 146.

Schulz, A., Dhar, S., Rylova, S., Dbaibo, G., Alroy, J., Hagel, C., Artacho, I., Kohlschütter, A., Lin, S., and Boustany, R.N., 2004. Impaired cell adhesion and apoptosis in a novel CLN9 Batten disease variant. *Ann Neurol.*, 56(3), 342-50.

Schulz, A., Mousallem, T., Venkataramani, M., Persaud-Sawin, D., Zucker, A., Luberto, C., Bielawska, A., Bielawski, J., Holthuis, J.C.M., Jazwinski, S.M., Kozhaya, L., Dbaibo, G.S., and Boustany, R.N., 2006. The CLN9 protein, a regulator of dihydroceramide synthase. *J Biol Chem.*, 281(5), 2784-94.

Settembre, C., Fraldi, A., Jahreiss, L., Spampanato, C., Venturi, C., Medina, D., de Pablo, R., Tacchetti, C., Rubinsztein, D.C., and Ballabio, A., 2008. A block of autophagy in lysosomal storage disorders. *Hum. Mol. Genet.*, 17(1), 119-129.

Sharp, J.D., Wheeler, R.B., Lake, B.D., Savukoski, M., Jarvela, I.E., Peltonen, L.,

Gardiner, R.M., and Williams, R.E., 1997. Loci for classical and a variant late infantile neuronal ceroid lipofuscinosis map to chromosomes 11p15 and 15q21-23. *Hum. Mol. Genet.*, 6(4), 591-595.

Shroyer, N. F., Lewis, R. A., Lupski, J. R., 2000. Complex inheritance of ABCR mutations in Stargardt disease: linkage disequilibrium, complex alleles, and pseudodominance, *Hum. Genet.*, 106(2), 244–248.

Siintola, E., Lehesjoki, A., and Mole, S.E., 2006a. Molecular genetics of the NCLs -- status and perspectives. *Biochimica Et Biophysica Acta*, 1762(10), 857-64.

Siintola, E., Partanen, S., Strömme, P., Haapanen, A., Haltia, M., Maehlen, J., Lehesjoki, A., and Tyynelä, J., 2006b. Cathepsin D deficiency underlies congenital human neuronal ceroid-lipofuscinosis. *Brain: A Journal of Neurology*, 129(Pt 6), 1438-45.

Siintola, E., Topcu, M., Aula, N., Lohi, H., Minassian, B.A., Paterson, A.D., Liu, X., Wilson, C., Lahtinen, U., Anttonen, A., and Lehesjoki, A., 2007. The novel neuronal ceroid lipofuscinosis gene MFSD8 encodes a putative lysosomal transporter. *Am J Hum Genet.*, 81(1), 136-46.

Skoog, W. A., Beck W. S., 1956. Studies on the fibrinogen dextran and phytohemagglutinin methods of isolating leukocytes. *Blood*, 11, 436-454.

Sleat, D., Zheng, H., and Lobel, P., 2007. The human urine mannose 6-phosphate glycoproteome. *Biochimica et biophysica acta*, 1774(3), 368–372.

Sleat, D.E., Ding, L., Wang, S., Zhao, C., Wang, Y., Xin, W., Zheng, H., Moore, D.F., Sims, K.B., and Lobel, P., 2009. Mass spectrometry-based protein profiling to determine the cause of lysosomal storage diseases of unknown etiology. *Molecular & Cellular Proteomics: MCP*, 8(7), 1708-1718.

Sleat, D.E., Sohar, I., Pullarkat, P.S., Lobel, P., and Pullarkat, R.K., 1998. Specific alterations in levels of mannose 6-phosphorylated glycoproteins in different neuronal ceroid lipofuscinoses. *The Biochemical Journal*, 334(Pt 3), 547-551.

Sleat, D.E., Donnelly, R.J., Lackland, H., Liu, C.G., Sohar, I., Pullarkat, R.K., and Lobel, P., 1997. Association of mutations in a lysosomal protein with classical late-infantile neuronal ceroid lipofuscinosis. *Science (New York, N.Y.)*, 277(5333), 1802-5.

Sleat, D.E., El-Banna, M., Sohar, I., Kim, K., Dobrenis, K., Walkley, S.U., and Lobel, P., 2008. Residual levels of tripeptidyl-peptidase I activity dramatically ameliorate disease in late-infantile neuronal ceroid lipofuscinosis. *Molecular Genetics and Metabolism*, 94(2), 222-33.

Sleat, D.E., Gin, R.M., Sohar, I., Wisniewski, K., Sklower-Brooks, S., Pullarkat, R.K., Palmer, D.N., Lerner, T.J., Boustany, R.N., Uldall, P., Siakotos, A.N., Donnelly, R.J., and Lobel, P., 1999. Mutational analysis of the defective protease in classic late-infantile neuronal ceroid lipofuscinosis, a neurodegenerative lysosomal storage disorder. *American Journal of Human Genetics*, 64(6), 1511-23.

Sleat, D.E., Lackland, H., Wang, Y., Sohar, I., Xiao, G., Li, H., and Lobel, P., 2005. The human brain mannose 6-phosphate glycoproteome: a complex mixture composed of multiple isoforms of many soluble lysosomal proteins. *Proteomics*, 5(6), 1520-

32.

Sleat, D.E., Wang, Y., Sohar, I., Lackland, H., Li, Y., Li, H., Zheng, H., and Lobel, P., 2006a. Identification and validation of mannose 6-phosphate glycoproteins in human plasma reveal a wide range of lysosomal and non-lysosomal proteins. *Molecular & Cellular Proteomics: MCP*, 5(10), 1942-56.

Sleat, D.E., Zheng, H., Qian, M., and Lobel, P., 2006b. Identification of sites of mannose 6-phosphorylation on lysosomal proteins. *Molecular & Cellular Proteomics: MCP*, 5(4), 686-701.

Smith, C. W., Chu, T. T., Nadal-Ginard, B., 1993. Scanning and competition between AGs are involved in 3' splice site selection in mammalian introns. *Mol. Cell. Biol.,* 13, 4939–4952.

Sondhi, D., Peterson, D.A., Giannaris, E.L., Sanders, C.T., Mendez, B.S., De, B., Rostkowski, A.B., Blanchard, B., Bjugstad, K., Sladek, J.R., Redmond, D.E., Leopold, P.L., Kaminsky, S.M., Hackett, N.R., and Crystal, R.G., 2005. AAV2-mediated CLN2 gene transfer to rodent and non-human primate brain results in long-term TPP-I expression compatible with therapy for LINCL. *Gene Therapy*, 12(22), 1618-32.

Steinfeld, R., Heim, P., von Gregory, H., Meyer, K., Ullrich, K., Goebel, H.H., and Kohlschütter, A., 2002. Late infantile neuronal ceroid lipofuscinosis: quantitative description of the clinical course in patients with CLN2 mutations. *American Journal of Medical Genetics*, 112(4), 347-54.

Steinfeld, R., Steinke, H., Isbrandt, D., Kohlschütter, A., and Gärtner, J., 2004. Mutations in classical late infantile neuronal ceroid lipofuscinosis disrupt transport of tripeptidyl-peptidase I to lysosomes. *Human Molecular Genetics*, 13(20), 2483-91.

Tammen, I., Houweling, P.J., Frugier, T., Mitchell, N.L., Kay, G.W., Cavanagh, J.A.L., Cook, R.W., Raadsma, H.W., and Palmer, D.N., 2006. A missense mutation (c.184C>T) in ovine CLN6 causes neuronal ceroid lipofuscinosis in Merino sheep whereas affected South Hampshire sheep have reduced levels of CLN6 mRNA. *Biochimica Et Biophysica Acta*, 1762(10), 898-905.

Tang, J., and Wong, R.N., 1987. Evolution in the structure and function of aspartic proteases. *Journal of Cellular Biochemistry*, 33(1), 53-63.

Teixeira, C., Guimarães, A., Bessa, C., Ferreira, M.J., Lopes, L., Pinto, E., Pinto, R., Boustany, R.N., Sá Miranda, M.C., and Ribeiro, M.G., 2003. Clinicopathological and molecular characterization of neuronal ceroid lipofuscinosis in the Portuguese population. *Journal of Neurology*, 250(6), 661-7.

Teixeira, C. A. F., S. Lin, M. Mangas, R. Quinta, C. J. P. Bessa, C. Ferreira, M. C. Sá Miranda, R-M. N. Boustany, and M. G. Ribeiro. 2006. Gene expression profiling in vLINCL CLN6-deficient fibroblasts: Insights into pathobiology. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease,* 1762(7), 637-646.

Tian, Y., Sohar, I., Taylor, J.W., and Lobel, P., 2006. Determination of the Substrate Specificity of Tripeptidyl-peptidase I Using Combinatorial Peptide Libraries and Development of Improved Fluorogenic Substrates. *J. Biol. Chem.*, 281(10), 6559-6572.

Tyynelä, J., Suopanki, J., Santavuori, P., Baumann, M., and Haltia, M., 1997. Variant late infantile neuronal ceroid-lipofuscinosis: pathology and biochemistry. *Journal of Neuropathology and Experimental Neurology*, 56(4), 369-75.

Tyynelä, J., Palmer, D.N., Baumann, M., and Haltia, M., 1993. Storage of saposins A and D in infantile neuronal ceroid-lipofuscinosis. *FEBS Letters*, 330(1), 8-12.

Tyynelä, J., Sohar, I., Sleat, D.E., Gin, R.M., Donnelly, R.J., Baumann, M., Haltia, M., and Lobel, P., 2000. A mutation in the ovine cathepsin D gene causes a congenital lysosomal storage disease with profound neurodegeneration. *The EMBO Journal*, 19(12), 2786-92.

Tyynelä, J., Suopanki, J., Baumann, M., and Haltia, M., 1997. Sphingolipid activator proteins (SAPs) in neuronal ceroid lipofuscinoses (NCL). *Neuropediatrics*, 28(1), 49-52.

Uusi-Rauva, K., Luiro, K., Tanhuanpää, K., Kopra, O., Martín-Vasallo, P., Kyttälä, A., and Jalanko, A., 2008. Novel interactions of CLN3 protein link Batten disease to dysregulation of fodrin-Na+, K+ ATPase complex. *Experimental Cell Research*, 314(15), 2895-905.

van Diggelen, O. P., Keulemans, J. L. M., Winchester, B., Hofman, I. L., Vanhanen, S. L., Santavuori, P., Voznyi, Y. V., 1999. A rapid Xuorogenic palmitoyl-protein thioesterase assay: Pre- and postnatal diagnosis of INCL, *Mol. Genet. Metab*., 66, 240–244.

Varki, A., and Kornfeld, S., 1980. Identification of a rat liver alpha-N-acetylglucosaminyl phosphodiesterase capable of removing "blocking" alpha-N-acetylglucosamine residues from phosphorylated high mannose oligosaccharides of lysosomal enzymes. *The Journal of Biological Chemistry*, 255(18), 8398-401.

Verkruyse, L.A., and Hofmann, S.L., 1996. Lysosomal Targeting of Palmitoyl-protein Thioesterase. *J. Biol. Chem.*, 271(26), 15831-15836.

Vesa, J., Chin, M.H., Oelgeschlager, K., Isosomppi, J., DellAngelica, E.C., Jalanko, A., and Peltonen, L., 2002. Neuronal Ceroid Lipofuscinoses Are Connected at Molecular Level: Interaction of CLN5 Protein with CLN2 and CLN3. *Molecular Biology of the Cell*, 13(7), 2410-2420.

Vesa, J., Hellsten, E., Verkruyse, L.A., Camp, L.A., Rapola, J., Santavuori, P., Hofmann, S.L., and Peltonen, L., 1995. Mutations in the palmitoyl protein thioesterase gene causing infantile neuronal ceroid lipofuscinosis. *Nature*, 376(6541), 584-7.

Vines, D.J., and Warburton, M.J., 1999. Classical late infantile neuronal ceroid lipofuscinosis fibroblasts are deficient in lysosomal tripeptidyl peptidase I. *FEBS Letters*, 443(2), 131-5.

von Figura, K. 1991. Molecular recognition and targeting of lysosomal proteins. *Current Opinion in Cell Biology,* 3(4), 642-6.

von Heijne, G., 1986. A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res*., 14, 4683.

von Heijne, G., 1987. Sequence Analysis in Molecular Biology: Treasure Trove or Trivial Pursuit. *Acad. Press*, 113-117.

Waheed, A., Hasilik, A., and von Figura, K., 1981. Processing of the phosphorylated recognition marker in lysosomal enzymes. Characterization and partial purification of a microsomal alpha-N-acetylglucosaminyl phosphodiesterase. *The Journal of Biological Chemistry*, 256(11), 5717-21.

Walus, M., Kida, E., Wisniewski, K.E., and Golabek, A.A., 2005. Ser475, Glu272, Asp276, Asp327, and Asp360 are involved in catalytic activity of human tripeptidyl-peptidase I. *FEBS Letters*, 579(6), 1383-1388.

Warburton, M.J., and Bernardini, F., 2002. Tripeptidyl peptidase-I is essential for the degradation of sulphated cholecystokinin-8 (CCK-8S) by mouse brain lysosomes. *Neuroscience Letters*, 331(2), 99-102.

Ward, D.M., Pevsner, J., Scullion, M.A., Vaughn, M., and Kaplan, J., 2000. Syntaxin 7 and VAMP-7 are soluble N-ethylmaleimide-sensitive factor attachment protein receptors required for late endosome-lysosome and homotypic lysosome fusion in alveolar macrophages. *Molecular Biology of the Cell*, 11(7), 2327-33.

Wheeler, R.B., Sharp, J.D., Schultz, R.A., Joslin, J.M., Williams, R.E., and Mole, S.E., 2002. The Gene Mutated in Variant Late-Infantile Neuronal Ceroid Lipofuscinosis (CLN6) and in nclf Mutant Mice Encodes a Novel Predicted Transmembrane Protein. *American Journal of Human Genetics*, 70(2), 537–542.

Winter, E., and Ponting, C.P., 2002. TRAM, LAG1 and CLN8: members of a novel family of lipid-sensing domains? *Trends in Biochemical Sciences*, 27(8), 381-3.

Wisniewski, K.E., Kida, E., Patxot, O.F., and Connell, F., 1992. Variability in the clinical and pathological findings in the neuronal ceroid lipofuscinoses: Review of data and observations. *American Journal of Medical Genetics*, 42(4), 525-532.

Wisniewski, K.E., Rapin, I., and Heaney-Kieras, J., 1988. Clinico-pathological variability in the childhood neuronal ceroid-lipofuscinoses and new observations on glycoprotein abnormalities. *American Journal of Medical Genetics. Supplement*, 5, 27-46.

Wlodawer, A., Durell, S.R., Li, M., Oyama, H., Oda, K., and Dunn, B.M., 2003. A model of tripeptidyl-peptidase I (CLN2), a ubiquitous and highly conserved member of the sedolisin family of serine-carboxyl peptidases. *BMC Structural Biology*, 3, 8.

Worgall, S., Sondhi, D., Hackett, N.R., Kosofsky, B., Kekatpure, M.V., Neyzi, N., Dyke, J.P., Ballon, D., Heier, L., Greenwald, B.M., Christos, P., Mazumdar, M., Souweidane, M.M., Kaplitt, M.G., and Crystal, R.G., 2008. Treatment of late infantile neuronal ceroid lipofuscinosis by CNS administration of a serotype 2 adeno-associated virus expressing CLN2 cDNA. *Human Gene Therapy*, 19(5), 463-74.

Wujek, P., Kida, E., Walus, M., Wisniewski, K.E., and Golabek, A.A., 2004. N-Glycosylation Is Crucial for Folding, Trafficking, and Stability of Human Tripeptidyl-peptidase I. *J. Biol. Chem.*, 279(13), 12827-12839.

applications to RNA splicing signals. *Journal of Computational Biology: A Journal of Computational Molecular Cell Biology*, 11(2-3), 377-94.

Zeman, W., 1976. The neuronal ceroid-lipofuscinosis. *Progress in Neuropathology*. Grune and Stratton, 3, 203-223.

Zeman, W., and Dyken, P., 1969. Neuronal ceroid-lipofuscinosis (Batten's disease): relationship to amaurotic family idiocy? *Pediatrics*, 44(4), 570-83.

Zhong, N., Moroziewicz, D.N., Ju, W., Jurkiewicz, A., Johnston, L., Wisniewski, K.E., and Brown, W.T., 2000. Heterogeneity of late-infantile neuronal ceroid lipofuscinosis. *Genetics in Medicine: Official Journal of the American College of Medical Genetics*, 2(6), 312-8.

12 Online sources

Crystal, R. 2007. Safety Study of a Gene Transfer Vector for Children With Late Infantilie neuronal Ceroid Lipofuscinosis. Weill Medical College of Cornell University, Foundation NsB (2007).
http://clinicaltrials.gov/show/NCT00151216

NICHD, 2007. Cystagon to Treat Infantile Neuronal Ceroid Lipofuscinosis.
http://clinicaltrials.gov/ct2/show/NCT00028262

NCL Mutation Database
http://www.ucl.ac.uk/ncl/mutation.shtml

Orchard, P., 2007. Stem Cell Transplant for Inborn Errors of Metabolism. Available at: http://clinicaltrials.gov/show/NCT00176904

StemCells, Inc. 2006. Study of the Safety and Preliminary Effectiveness of Human Central Nervous System (CNS) Stem Cells (HuCNS-SC) in Patients With Infantile or Late Infantile Neuronal Ceroid Lipofuscinosis (NCL).
http://clinicaltrials.gov/ct2/show/NCT00337636

13 Acknowledgments

The work presented in this thesis began in late 2003. Since that time all the experiences I went through and knowledge I have gained would not have been possible without the help of a considerable number of people, to whom I feel truly grateful. There are those who have accompanied and supported me throughout all this time, and those who at a particular time point contributed to the success of my work… to everyone my sincere thank you.

Financial support was provided by Fundação para a Ciência e a Tecnologia and the Vaincre les Maladies Lysosomales association, which are greatly acknowledged, as well as Instituto de Ciências Biomedicas Abel Salazar, from Universidade do Porto, my host academic institution.

In the course of this project I had the opportunity to perform my work at several institutes and laboratories. I wish to acknowledge the respective directors and department leaders for providing the facilities and resources necessary for my research.

I spent my first year as a PhD student at the Lysosome and Peroxisome Biology Unit at IBMC and wish to thank group leader Dr. Clara Sá Miranda for providing the conditions that enabled the start of my PhD project. I am also grateful to my former lab colleagues Andrea, Daniel, Andreia, Manuel, Claudia e João for all the fun times and good atmosphere at the lab.

In 2005 I had the opportunity to work for a brief summer at Professor's Richard Pagano laboratory at the Mayo Clinic in Rochester (MN, USA). I wish to acknowledge him and his lab members for a very warmth welcome, and thank technician Christine Wheatley who was always present to help me at the lab and took me under her wing during the time I was there. I must thank Professor Pagano and is team for their expertise in microscopy and imaging, which has become one of my favorite technical fields. And last, Travis and Giovanni for making my stay in Rochester so pleasant and joyful.

The majority of the thesis work was done at the Enzimologia unit at the Centro de Genética Médica Jacinto Magalhães (INSA). I would like to acknowledge unit leader Dr. Lucia Lacerda for providing the resources that I needed to accomplish my work. I wish to express my gratitude to all Enzimologia members for their companionship and support: Isaura, Eugénia, Elizabete, Sónia, Lurdes, Rita, Célia, Helena, Dr. Olga Amaral and Dr. Manuela Lemos. And last but not least the guys, Francisco and Aureliano. It has been great to work you all. I would also like to give a big thanks to Clementina, who is always

ready to help with any bureaucratic problem, and Ana and Rosa, without whom nothing would get done at the lab.

The final period of my PhD project was spent in Helsinki, Finland. I wish to acknowledge Professors Anu Jalanku and Aija Kyttälä for their guidance and support during my stay. I had a tight timeframe and busy schedule and your determination was invaluable to my success. I also want to give a big thanks to Mia Schmiedt, with whom I worked closely every day. It was a pleasure, thank you for you friendship and for making those days so joyful. I also want to express my gratitude to all lab members for being so nice to a visitor, Auli, Jenni, Annina, Mari, Terhi, Kaija. I brought nothing but good memories from your country.

And let's not forget all my colleagues from “mokifo” (you know what I mean). Sandra, Natália, Liliana, Francisca, thank you for all the great times. Mia, I hope I got to pass something positive to you during the time we've spent working together, I'm sure I've learned from your company. Carla, I'm very grateful to have had the chance to work with you. You've been there from the start, and have taught me so much. Thank you for your friendship.

I wish to acknowledge my co-supervisor, Professor Jorge Azevedo, to whom I always felt free to go for guidance and advice with regard to my work. Thank you for your esteemed contributions.

And finally, I would like to express my gratitude to my supervisor, Dr. Maria Gil Ribeiro, who's positive attitude and enthusiasm towards science kept me motivated throughout the project. Without you dedication, and exceptional guidance I would not be here today. Thank you for trusting my capabilities and for all I've learned throughout this time.

And now to those that, although not quite getting what I kept talking about for so long, never stopped listening... my family and friends.

I am very grateful to my parents, who always incited me to pursue my education and academic studies, believing in my potential. Thank you for helping me become who I am today. And thanks for all those books in the attic that made my night time reading, seems like were of use after all. I also want to congratulate my brother Paulo and his wife Catarina for their great job on giving me those lovely nephews that inspired me during the time I spent writing this thesis.

I want to give a big thanks to my close friends, Pedro, Joana, Valter, Cris, Jorge,

Mário and Gi, who's companionship kept me up during the downturns of every scientific study, and got me in close touch with reality. Thank you for all the good times.... now just keep them coming.

Finally, I want to express my deepest gratitude to my loving wife Patricia. Thank you for all the love and your never ending patience. You're the one who's kept me going all this time.

…....and to everyone I didn't mention, I did not forget about you... I just run out of space!

This thesis is my small contribution to the world of science, in hope that diluted in the immense pool of the scientific community we as a whole can move forward.

"What is a scientist after all? It is a curious man looking through a keyhole, the keyhole of nature, trying to know what's going on."

Jacques Yves Cousteau

14 Appendix

14.1 Publication III (manuscript)

The neuronal ceroid lipofuscinosis protein CLN5: New insights into cellular maturation, transport and consequences of mutations

Mia-Lisa Schmiedt[1*], Carlos Bessa[2,3*], Claudia Heine[1], Maria Gil Ribeiro[2], Anu Jalanko[1], and Aija Kyttälä[1$]

[1]National Institute for Health and Welfare (THL), Public Health Genomics Unit, Helsinki, Finland and FIMM, Institute for Molecular Medicine in Finland,
[2]Instituto Nacional de Saúde Dr. Ricardo Jorge (INSA), Unidade de Investigação e Desenvolvimento, Departamento de Genética, Porto, Portugal,
[3]Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto

[*] These authors contributed equally to the work

$ Address correspondence to: Aija Kyttälä, National Institute for Health and Welfare (THL), Public Health Genomics Unit, Biomedicum P.O.Box 104, FIN-00251 Helsinki, Finland. Tel. +358-20 6108976, Fax. +358-20 610 8960 Email: aija.kyttala@thl.fi

ABSTRACT

Neuronal ceroid lipofuscinoses (NCLs) represent a group of children's inherited neurodegenerative disorders caused by mutations in at least eight different genes. Mutations in the *CLN5* gene result in the Finnish variant late infantile NCL characterized by gradual loss of vision, epileptic seizures and mental deterioration. The *CLN5* gene encodes a lysosomal glycoprotein with unidentified function. In this study, we have -used both transient and stable expression systems for the characterization of CLN5 focusing on the localization, processing and intracellular trafficking. We show that CLN5 is proteolytically cleaved and the mature polypeptide is transported to the lysosomes. Our data provide the first evidence that soluble CLN5 protein can also undergo mannose-6-phosphate receptor-independent trafficking to the lysosomes. We studied the localization and maturation of the CLN5 carrying the previously uncharacterized vLINCL disease causing mutations in HeLa cells. All analyzed disease mutations disturb the lysosomal trafficking of the mutated CLN5 proteins. The level of lysosomal targeting does not however correlate to disease onset, indicating that CLN5 may also function outside lysosomes. This study furthers understanding of the basic properties of the CLN5 protein, necessary for the characterization of the consequences of disease mutations and for the planning of future therapies for vLINCL.

KEY WORDS

Neuronal ceroid lipofuscinoses, NCL, CLN5, lysosomal trafficking, protein N-glycosylation

INTRODUCTION

Neuronal ceroid lipofuscinoses (NCLs) comprise a group of progressive neurodegenerative disorders of childhood. The estimated incidence is 1:12500 in the USA and Nordic countries and approximately 1:100 000 worldwide (Haltia, 2006; Santavuori, 1988). NCL diseases with characterized gene defects are recessively inherited and to date, eight genes are known to underlie this group of disorders (Jalanko and Braulke, 2008). Despite the genetic heterogeneity, clinical and neuropathological findings are highly similar and suggest common molecular mechanisms behind the NCL disorders. Although the first gene defects were described more than ten years ago, the functions of most NCL proteins has remained unclear (Kyttala, et al., 2006).

Mutations in the *CLN5* gene (MIM# 608102) result in a variant late infantile form of NCL, enriched in the Finnish population. To date, 20 disease-causing mutations have been reported (http://www.ucl.ac.uk/ncl). The *CLN5* gene is located in chromosome 13q21–q32 and encodes a 407 amino acid polypeptide with a predicted molecular mass of 46 kDa. It lacks homology to any known protein and its function is unclear (Savukoski, et al., 1998). Interestingly, human CLN5 contains four initiator methionines and previous *in vitro* translation analysis of the *CLN5* cDNA demonstrated the synthesis of four polypeptides with apparent molecular weights from 39 to 47 kDa (Isosomppi, et al., 2002). Due to eight potential *N*-glycosylation sites, the CLN5 protein has been shown to have an apparent molecular mass of 52–75 kDa with both high-mannose-type as well as complex-type sugars (Isosomppi, et al., 2002; Vesa, et al., 2002). The solubility of the CLN5 protein has been controversially discussed (Bessa, et al., 2006; Isosomppi, et al., 2002; Savukoski, et al., 1998; Vesa, et al., 2002). However, it has been reported that human CLN5 contains mannose 6-phosphate (Man 6-P) residues on high-mannose type oligosaccharides linked to Asn320, Asn330, and Asn401, supporting the existence of soluble CLN5 variants (Sleat, et al., 2006). Furthermore, the mouse Cln5 protein has been reported to be soluble (Holmberg, et al., 2004). In overexpression studies, CLN5 has been located to the ER (Vesa, et al., 2002) and to lysosomes (Isosomppi, et al., 2002; Vesa, et al., 2002). In addition, CLN5 has been found in axons of neuronal cells (Holmberg, et al., 2004). Co-immunoprecipitation and *in vitro* binding assays have shown that CLN5 interacts with two other NCL proteins, CLN2 (TPP1) and CLN3 (Vesa, et al., 2002). However, some of the basic properties of the CLN5 protein have not yet been elucidated. Here we have examined the localization, maturation and intracellular trafficking of CLN5. Consequently, the impact of several disease causing mutations were analyzed in light of our novel data.

MATERIALS AND METHODS

Recombinant *CLN5*-cDNA constructs

Human wild type CLN5-pCMV (aa 1-407) and CLN5-pCMV carrying the different disease causing mutations p. Y392X (Fin_{Major}), p. E253X (SWE) and p. D279N (EUR), have been described previously (Isosomppi, et al., 2002; Vesa, et al., 2002). CLN5 carrying the p. R112P (POR) and p. R112H (COL) mutations were generated into CLN5-pCMV (human CLN5: GenBank NM_006493.2) construct by site-directed *in vitro* mutagenesis, using the QuickChange mutagenesis kit from Stratagene, following the manufacturer's instructions**.** All used expression constructs of mutated CLN5 (p. R112H, p. R112P, p. D279N, p. E253X, p. Y392X) start from the first methionine M1. GFP-CLN5 constructs starting from methionines M1 to M4 were cloned in frame with an *N*-terminal GFP-tag in vector pEGFP-C1 (Clonetech). Full-length mouse Cln5 cDNA (aa 1-341) with a *C*-terminal myc/his-tag, was generated by cloning the Cln5cDNA into pcDNA3.1A/*Myc-His* expression vector (Invitrogen). The construction of CLN3-pCMV was described earlier (Jarvela, et al., 1998). All constructs were verified by sequencing.

Cell culture and transfection

HeLa and COS-1 cells were cultured in Dulbecco´s modified Eagle´s medium (DMEM), supplemented with 10 % fetal calf serum (FCS) and 1 % antibiotics (penicillin/streptomycin). Human neuroblastoma cells SH-SY5Y cells were maintained 1:1 DMEM and Ham´s F12 medium (Sigma), supplemented with 10 % FCS,

penicillin/streptomycin and 0.1 % non-essential amino acids. Transient transfections were performed on 6-well plates with Fugene HD (Roche Diagnostics) transfection reagent, following the manufacturer's instructions. For localization studies of human CLN5 a stable cell line was established (CLN5-SH-SY5Y). Human neuroblastoma cells (SH-SY5Y) were co-transfected with human CLN5-pCMV5 and pREP4 vector (Invitrogen). The latter vector conferred the resistance to hygromycin B (200 µg/ml) (InvivoGen), which was used for selection of the stably transfected cells and thereafter maintained in the culture media at the concentration of 50 µg/ml. Mannose-6-phosphate-receptor double deficient mouse fibroblasts ($MPR^{-/-}$; MPR46 and MPR300 deficient) were kindly provided by Thomas Braulke (Hamburg, Germany) (Pohlmann, et al., 1995). The mouse fibroblasts were cultured in DMEM, supplemented with 20 % FCS and 1% antibiotics (penicillin/streptomycin). Transient transfection of the $MPR^{-/-}$ cells in the 6-well plates was performed by using the Lipofectamine 2000 transfection reagent (Invitrogen), according to the manufacturer's instructions.

Antibodies

A novel rabbit polyclonal antibody (CLN5-C/32) against a synthetic peptide (EEIPLPIRNKTLSGL) based on the C-terminus extremity of the human CLN5 was developed by Eurogentec (Germany). The specificity was tested in Western Blot and immunfluorescence analysis using a stable cell line CLN5-SHSY-5Y (described above) and SH-SY5Ycells as control. The CLN5-C/32 antibody (1:300-1:500) was then used in all Western Blot analyzes to detect CLN5. In immunofluorescence studies, CLN5 was detected either by a novel CLN5-C/32 antibody, by a previously described rabbit antibody (1RmI-4) raised against the GST-Cln5 (aa 40-284) or by a guinea pig antibody (1GmII-3) raised against the GST-Cln5 (aa 40-341) (Holmberg, et al., 2004). CLN3 was detected with the peptide antibody 385 (Luiro, et al., 2001).The mouse monoclonal anti-GFP antibody was obtained from Santa Cruz Biotechnology. LAMP-1 specific antibodies made against human (H4A3) and mouse (1D4B) antigens were obtained from the Developmental Studies Hybridoma Bank developed under the auspices of the NICHD, National Institutes of Health, and maintained by the University of Iowa (Iowa City). Mouse monoclonal PDI (Protein Disulfide Isomerase) antibody (1D3) was obtained from Stressgen. Mouse monoclonal c-Myc antibody (9E10) was purchased from Santa Cruz Biotechnology (Santa Cruz, CA).

Immunfluorescence staining

For immunofluorescence analysis, transiently or stably transfected cells were grown on coverslips and fixed 48 h after transfection in ice cold methanol for 3-5 min. The primary antibodies were diluted in 0.5 % BSA in PBS and incubated for 1 h at 37 °C. The secondary antibodies were also diluted in 0.5 % BSA in PBS and incubated for 1 h at room temperature. All secondary antibodies used were from Jackson ImmunoResearch. The labelled coverslips were mounted using GelMount (Biomeda Corp.) and visualized using Leica DMR confocal microscope with TCS NT software (Leica Microscope and Scientific Instruments Group). ImageJ, Adobe Photoshop and Adobe Illustrator software were used for image processing. Brightness and contrast were adjusted independently for each channel.

Western Blotting

Cells were collected and lysed for 20 minutes on ice (50 mM tris-HCl pH 7.5, 300 mM NaCl, 1 % TritonX-100, protein inhibitor cocktail (Roche Diagnostics)) followed by a 10 minute centrifugation at 13200 rpm. After heat denaturation with Laemmli buffer containing 2 % β-mercaptoethanol, samples were loaded and separated on 10-11 % SDS-PAGE gels. Proteins were transferred to nitrocellulose membranes (Amersham Biosciences) by electroblotting and Ponceau protein stain was used to confirm protein loading and even transfer. Skimmed milk powder (5 %) in TBS-T was used for blocking and as dilution buffer for primary antibodies. Secondary HRP-conjugated antibodies were obtained from DAKO and were used according to the manufacturer's instructions. The ECL Western Blotting Detection Kit (Amersham) was used as HRP substrate for the detection of antigens.

Pulse-chase analysis and immunoprecipitation

For immunoprecipitation analysis CLN5-SH-SY5Y cells were washed twice with PBS and metabolically labeled by starving them in methionine-free DMEM for 30 min and thereafter labeling with 100 µCi/500 µl of [^{35}S] methionine (Perkin-Elmer) for 1 h. After a washing step with PBS, the cells were chased in regular DMEM and Ham´s F12 medium (1:1 mix), supplemented with 10 % FCS, penicillin/streptomycin and 0.1 % non-essential amino acids for the indicated times. Cells were pelleted by centrifugation and lysed (see Western Blotting). Lysed cells were immunoprecipitated with CLN5-C/32 antibody (1:100) and A/G sepharose beads (Santa Cruz). The samples were loaded and separated on a 10 % SDS-PAGE gel. The gel was fixed with 10 % acetic acid/25 % Isopropanol and the radioactive signal was intensified with 0.7 M sodium salicylate. The dried gel was exposed to autoradiography film (Amersham) for 1 day.

Analysis of protein processing and glycosylation

Stably transfected SH-SY5Y cells (CLN5-SH-SY5Y) were grown to 80 % confluence on 6-well plates before treatments with different reagents. Reagents were diluted in 1:1 DMEM and Ham´s F12 medium (Sigma), supplemented with 10 % FCS, penicillin/streptomycin and 0.1 % non-essential amino acids. HeLa cells were treated 48 h post transfection.

Blocking of protein synthesis was performed by treating the cells with 50 µg/ml cycloheximide (Sigma-Aldrich) up to 24 hours. To inhibit anterograde trafficking between the ER and Golgi apparatus, the cells were treated with 2.5 µg/ml of brefeldin A (Epicentre Biotechnologies) for 5 h. Protein N-glycosylation was prevented by the addition of 10 µg/ml tunicamycin (MP Biomedicals) for 6 and 24 hours. Tunicamycin blocks the synthesis of dolichol pyrophosphate-N-acetylglucosamine, a key enzyme in glycoprotein biosynthesis in the ER. Cell pellets of SH-SY5Y cells (stable CLN5 and untransfected) or HeLa cells were lysed and the protein concentration was equalised after measuring the amount of protein with BCA Assay protein Quantification kit (Uptima). Samples were separated by SDS-PAGE, immunoblotted and probed for CLN5.

Enzyme digestion with peptide N-glycosidase F (PNGase F) and endoglycosidase H (Endo H)

Cell lysates of CLN5-SH-SY5Y and transfected COS-1 cells were treated with Endo H (New England BioLabs) and PNGase F (Roche) for glycosylation studies. 20 µg of the total protein was mixed with 1 µL of 10x Glycoprotein Denaturating Buffer and water to make a 10 µl total reaction volume. Glycoproteins were then denatured by heating reaction at 100 °C for 10 min. Endo H treatment: 2 µl 10x G5 reaction buffer, 5-7 µl H_2O and 1-3 µl Endo H were added and incubated at 37 °C for 24 h. PNGase F treatment: 2 µl 10x G7 reaction buffer, 2 µl NP-40, 3 µl H_2O and 3 µl PNGase F were added and incubated at 37 °C for 24 h. Separation of digested proteins were visualized by SDS-PAGE, immunoblotting and finally CLN5 probing.

RESULTS

Proteolytic cleavage of CLN5 produces a single polypeptide from all four initiation methionines

The human *CLN5* gene encodes for a protein of 407 amino acids with a calculated molecular weight of about 46 kDa (Savukoski, et al., 1998). Its sequence possesses four in-frame AUG codons (Methionine M 1-4), at amino acid positions 1, 30, 50 and 62, that have been shown to be used *in vitro* but also by COS-1 cells (Isosomppi, et al., 2002; Vesa, et al., 2002). SignalP is a software prediction model that aims to determine the existence and position of signal peptide cleavage sites in proteins (Bendtsen, et al., 2004). Analysis of CLN5 sequences consistently suggest the presence of such a cleavage site in the CLN5 polypeptides starting from M4, M3 and M2, even though with different probabilities, at a position corresponding to amino acid 96 of the "full" CLN5 sequence (NP_006484, NCBI database). However, the likelihood of the existence of a signal peptide cleavage site in M1 is very low (Supp. Table S1). This is most likely due to the distance of the signal peptide region to the N-terminus extremity. In fact, the

SigCleave prediction model, which does not make use of a cut-off to disregard internal sites (von Heijne, 1986), attributes similar scores for all M1-4 CLN5 sequences.

To further address the usage of alternative initiation methionines by the cell machinery, several constructs were produced based on the human *CLN5* cDNA. A GFP tag was introduced at the N-terminus of each of the initiation methionines (Fig. 1A), and the constructs were used to transiently transfect COS-1 and HeLa cells. To obtain additional tools for the detection of the CLN5 protein, we produced a new peptide antibody, CLN5-C/32 against the C-terminal region (aa 393 - 407) of human CLN5. This novel antibody is specific for the detection of overexpressed CLN5 in Western Blot, in immunoprecipitation and immunofluorescence analyzes.
COS-1 cell lysates were analyzed by Western blotting using both the novel C-terminal human CLN5 antibody and the GFP-specific antibody to detect both the N- and C-terminal regions of the expressed GFP-CLN5 polypeptides. Detection with the CLN5 antibody showed a single 60 kDa polypeptide from constructs M1, M2, M3 and M4 (Fig. 1B) indicating that cleavage of the polypeptides starting from the different methionines had occurred at the same amino acid site. Correspondingly, different sizes of polypeptides were detected with the GFP antibody for the same constructs, approximately 43, 41, 39,5 and 38 kDa for M1, M2, M3 and M4 respectively (Fig. 1C). Some lower molecular weight polypeptides were also observed for each construct with the GFP antibody, but they vary between experiments, and as such are likely to be derived from protein degradation.

Localization of the overexpressed GFP-CLN5 constructs was also studied by immunofluorescence in COS-1 and HeLa cells. The transfected cells were immunostained with the novel C-terminal CLN5 specific antibody, together with the lysosomal marker LAMP-1 and analyzed using a confocal microscope. In accordance with the observations of the Western blot analyzes, signals of CLN5 and GFP were located in different compartments and only minimal co-localization was detected. This observation supports that the N- and the C-termini of the expressed proteins are separated. While the CLN5 specific antibody presented a vesicular pattern co-localising with the endosomal/lysosomal organelles in all M1-4 GFP-CLN5 constructs, GFP fluorescence showed a perinuclear reticular pattern resembling the ER (localization of GFP-CLN5 M2 is shown as an exemplary in Fig. 1D, data not shown for the other constructs). Expression of these constructs in HeLa cells, followed by co-staining with an ER-specific marker PDI, indicated that when the GFP was expressed in front of M1 to M4, the staining pattern often exceeded the ER-specific labelling suggesting that the N-terminal GFP have retained at the cytosolic site of the ER.

Altogether these data indicate that targeting of CLN5 to the lysosomes is preceded by an N-terminal cleavage that is not dependent on the position of a starting methionine. Therefore this cleavage step is most likely a signal peptide cleavage.

CLN5 is highly glycosylated and appears as a proform and a mature protein in stably transfected SH-SY5Y cells

Most NCL proteins are not detectable endogenously by immunofluorescence or Western Blot analyzes using the currently available antibodies. Therefore, it was necessary to set up a system with stable, but not excessive, overexpression of CLN5. Consequently, a stable human CLN5 expressing cell line (CLN5-SH-SY5Y) was generated. In Western Blot analysis of the CLN5-SH-SY5Y cell lysates, the novel C-terminal peptide antibody detected two CLN5 polypeptides of approximately 60 kDa and 50 kDa (Fig. 2A).

To investigate whether these bands represented the pro- and mature forms of CLN5 we analyzed the maturation of the protein after treatment of cells with cycloheximide (chx) to stop protein synthesis. The chx treatment of the CLN5-SH-SY5Y cells led to a marked decrease of the 60 kDa polypeptide after 6 hours. The 50 kDa polypeptide was clearly visible after 6 hours and still detectable after 24 hours of chx treatment, suggesting that the 50 kDa polypeptide represents the mature human CLN5 and that the 60 kDa polypeptide is a proform (Fig. 2A). This finding was verified by

metabolic labelling and pulse-chase experiments of CLN5-SH-SY5Y cells, which indicated that the 60 kDa proform was present immediately after 1h radioactive pulse and was fully maturated into the 50 kDa polypeptide after 4 hours of chase (Fig. 2B). The observation of the 60 kDa polypeptide immediately after the radioactive pulse suggests a co-translational processing which is in agreement with the hypothesis of N-terminal signal peptide cleavage.

Treatment of CLN5-SH-SY5Y cells with tunicamycin, which blocks the initiation of the protein N-glycosylation in the ER and therefore should only affect the 60 kDa proform of CLN5, further supported these findings. After 6 hours of tunicamycin treatment the intensity of the 60 kDa polypeptide was highly reduced and an unglycosylated 35 kDa band appeared. The 50 kDa polypeptide was still visible after a 6 h incubation period. After 24 hours of tunicamycin treatment signals for both, the pro- and mature form werebarely detectable (Fig. 2C).

The effects of protein glycosylation on CLN5 were also analyzed by digestion of CLN5-SH-SY5Y cell lysates with EndoH and PNGaseF. These enzymes remove N-linked carbohydrates from the glycosylated proteins but are not able to block the initiation of the glycosylation like tunicamycin. EndoH treatment, which removes only high mannose-type sugars, resulted in a downshift of both the 60 and 50 kDa bands. Interestingly, the downshift did not result in one band but in two distinct polypeptides at about 40 kDa and 37 kDa of size (Fig. 2D). Moreover, treatment with PNGaseF, which is also able to remove complex-type sugars, led to a further reduction of the 60 and 50 kDa polypeptides resulting in two distinct polypeptides at about 36 kDa and 35 kDa. The observation that both polypeptides were found to contain complex type sugars indicates that they have reached the Golgi complex (Fig. 2E). These findings are also well in line with the earlier published data (Isosomppi, et al., 2002; Vesa, et al., 2002).

We also analyzed how the proform and the mature CLN5 polypeptides were affected by a block in the intracellular transport. Therefore, we treated CLN5-SH-SY5Y cells with Brefeldin A (BFA) and immunoblotted the cell lysates. BFA blocks the anterograde transport from the ER to the Golgi apparatus, and leads to fusion of the Golgi compartment and its ingredients with the ER. However, anterograde trafficking from the trans-Golgi network (TGN) occurs normally. BFA treatment resulted in a reduction of the 60 kDa band to a strong signal at about 55 kDa, whereas the mature 50 kDa polypeptide was still visible after 5 hours of exposure (Fig. 2F). This observation is in line with our pulse chase experiments supporting that the 50 kDa polypeptide derives from the 60 kDa form at the trans-Golgi-network (TGN) or a post-Golgi location. Additional digestion after BFA treatment with EndoH or with PNGaseF did not result in different polypeptide sizes compared to the BFA-untreated controls, indicating that BFA treatment only resulted in changes in the glycosylation status of the 60 kDa proform of CLN5.

Overall these results demonstrate that human CLN5 is synthesized as a preproprotein; after co-translational removal of the N-terminal signal peptide, the protein is N-glycosylated with both high mannose and complex type sugars, and almost 50 % of the molecular size of the CLN5 protein is made of carbohydrates. Trimming of these carbohydrate structures during the intracellular trafficking of CLN5 is part of the maturation of the protein together with the processing step described above. Our data also indicate that CLN5 undergoes additional modifications.

Mannose 6-phosphate receptor mediated pathway is not crucial for the trafficking of CLN5 to the lysosomes

Most soluble lysosomal proteins contain mannose 6-phosphate (Man-6-P), a specific carbohydrate modification. This modified residue is recognised by mannose-6-phosphate-receptors (MPRs) that direct the targeting to the lysosomal compartment (Pohlmann, et al., 1995). It has been suggested that CLN5 uses the MPR-dependent pathway for its lysosomal trafficking (Kollmann, et al., 2005; Sleat, et al., 2006), but actual utilization of the MPR dependent pathway has not been shown experimentally.

To initially assess the trafficking of CLN5, we studied MPR deficient fibroblasts and analyzed the intracellular localization of mouse Cln5. MPR deficient mouse fibroblasts were transiently transfected with C-terminally myc-tagged mouse Cln5 construct and localization of the protein was analyzed by confocal microscopy. Interestingly, Cln5 was able to traffic to lysosomes in the absence of MPRs, and co-localized with the co-transfected lysosomal transmembrane protein CLN3. CLN3 and the lysosomal marker LAMP-1 were used here as controls for the MPR-independent trafficking, as they are transmembrane proteins and therefore do not utilize the MPR pathway (Fig. 3A-D).

This novel finding indicates that Cln5 is able to use an alternate MPR-independent route to lysosomes in addition to the fact that it has been found to contain three Man-6-P-tagged residues and therefore most probably also uses the classical MPR pathway (Sleat, et al., 2006).

Consequences of the vLINCL disease causing mutations on localization and maturation of the CLN5

To improve the understanding on the impact of different disease causing mutations on localization and maturation of the CLN5 protein and on the vLINCL disease phenotype, HeLa and COS-1 cells were transiently transfected with human CLN5 cDNA carrying the mutations in p.R112P, p.R112H, p.E253X, p.D279N, or p.Y392X. The p. D279N (EUR) and p.E253X (SWE) mutations have been previously described, but the intracellular localization of the latter has not been studied (Holmberg, et al., 2000; Savukoski, et al., 1998; Vesa, et al., 2002). Additionally, we analyzed the recently described, but also uncharacterized disease mutations hitting the amino acid 112, found in a Portuguese family (p.R112P; POR) and a Colombian patient (p.R112H; COL). The Portuguese patient is compound heterozygous carrying a null mutation (p.Q189X) in one allele and two mutations (p.R112P and p.D279N) in the other allele (Bessa, et al., 2006). The Colombian patient has been described to present a juvenile-onset phenotype (Pineda-Trujillo, et al., 2005). CLN5 carrying the most common vLINCL disease causing mutation, Fin_{Major} (p.Y392X) was also included in our studies, since there is controversy over the localization of this mutated polypeptide as discussed in Isosomppi, et al., 2002; Savukoski, et al., 1998; Vesa, et al., 2002.

Immunofluorescence analyzes showed that wt CLN5 localized to the lysosomes in HeLa cells, determined by co-localization with a lysosomal marker LAMP-1 (Fig. 4 A-C). All the studied mutations resulted in a different staining pattern compared to the wt, co-localising well with the ER marker PDI at steady state (Fig. 4D-L and Supp. Fig. S3). Since CLN5 carrying the Fin_{Major} or EUR mutations have been reported to be able to leave the ER and traffic to the Golgi complex (Isosomppi, et al., 2002) or even to the lysosomes (Vesa, et al., 2002), we also wanted to study whether the other mutations also allow the exit of the polypeptides from the ER. To analyze this, all the mutated CLN5 constructs were expressed in HeLa cells and the proteins were chased for 5 hours in the presence of chx. Staining with a CLN5-specific antibody showed that the polypeptides carrying the mutations hitting the amino acid 112 (p.R112P/H) were relatively stable but unable to exit the ER (Fig. 5A-F). The CLN5 protein carrying the EUR (p.D279N) mutation was found in the ER of 90 % of the transfected cells, but in about 10 % of the transfected cells the mutated polypeptide showed lysosomal localization (Fig. 5 G-I). The CLN5 carrying the SWE (p.E253X) or Fin_{Major} (p.Y392X) resulted in an unstable polypeptide not detectable after five hours of chx chase. Shorter incubation with chx (1h) indicated that most of the mutated polypeptides were still in the ER. Both SWE (p.E253X) (Fig. 5 J-L) and Fin_{Major} (p.Y392X) (Fig. 5 M-O) polypeptides were also detected in the vesicular structures, which however, did not co-localize with a lysosomal marker LAMP-1.

To analyze the consequences of the missense mutations on maturation of CLN5, COS-1 cells were transiently transfected with CLN5 carrying the different mutations. The cell lysates were then digested with EndoH or PNGaseF and the size of the mutated polypetides were analyzed by Western Blot using the CLN5 specific antibody (Fig. 6). As reported earlier, CLN5 carrying the EUR

mutation (p.D279N) was found to contain complex type sugars. This is also in line with our immunofluorecence studies that indicated some of the protein was able to reach the lysosomes. Interestingly, also a higher glycosylated form of 65 kDa was detected for the p.D279N polypeptide. In fact, this mutation results in an additional consensus sequence for N-glycosylation and our observations suggest that this novel glycosylation site is also utilized *in vivo*. Somehow unexpected, the polypeptides carrying R112 mutations (p.R112P/H) were also found to contain complex-type sugars although imunofluorescence analysis had suggested a stable ER localization. This may imply that CLN5 undergoes a retrograde transport step from the Golgi to the ER.

DISCUSSION

Since the initial characterization of the CLN5 protein underlying the Finnish variant late infantile NCL, both the localization and solubility of the protein have been under debate (Bessa, et al., 2006; Holmberg, et al., 2004; Isosomppi, et al., 2002; Savukoski, et al., 1998; Vesa, et al., 2002). Therefore, there was a clear necessity for a study further elucidating the basic characteristics of the protein. Here we used both transient and stable expression systems in different cell lines and established a novel C-terminal antibody for CLN5 to clarify the localization, maturation and trafficking of the protein. Consequently, the effects of the vLINCL disease causing mutations were examined in the light of the new data.

Human CLN5 was originally predicted to be a transmembrane protein of 407 amino acids (Savukoski, et al., 1998). Later, two contradictory studies on CLN5 solubility and localization were reported. Analysis of CLN5 with an N-terminal antibody (raised against aa 1-75) detected CLN5 in the supernatant only after Triton X-114 fractionation, and the protein was localized to the ER (Vesa, et al., 2002). However, with a peptide antibody raised against the amino acids 258-273 of CLN5, the protein was localized to lysosomes and found to be secreted, and therefore was reported to represent a soluble protein (Isosomppi, et al., 2002; Vesa, et al., 2002). Analysis of the mouse Cln5 protein later provided support for the notion of a soluble, lysosomal protein (Holmberg, et al., 2004; Vesa, et al., 2002) (all described CLN5 antibodies are listed in the Supp. Table S2).

We could experimentally demonstrate for the first time that CLN5 is N-terminally processed in the ER and the processing results in equal size polypetides independent of the position of the starting methionine in human CLN5. The observed proteolytic cleavage step of CLN5 most likely represents a co-translational signal sequence cleavage resulting in a pro-form of CLN5 which undergoes further maturation during it's trafficking to the lysosomes. The presented data therefore can explain the earlier discrepancies about the solubility of CLN5. Detection of the long form of the protein (starting from M1) with the N-terminal antibody (aa 1-75) showed only the non-cleaved, recombinant CLN5 protein whereas the peptide antibody raised against the central part of CLN5 detected the lysosomal mature protein (Isosomppi, et al., 2002; Vesa, et al., 2002). Recently, endogenous CLN5 polypeptides detected from human fibroblasts were reported to be membrane bound (Bessa, et al., 2006). However, using the same antibody in this study, we could not reproduce those results in our cell systems and concluded that either the antibody was not specific for CLN5 in Western blot analysis, or it could recognize yet unidentified splice variants, which is suggested for CLN5 (Savukoski, et al., 1998). It is noteworthy that some programs still predict two hydrophobic regions for the CLN5 protein e.g. PHDhtm (aa 74-91 and aa 356-375) and TMpred (aa 75-91 and aa 352-373). The first region most probably represents the predicted signal sequence, whereas the possible second hydrophobic region could facilitate the membrane association of the protein rather than be a real transmembrane domain *per se.*

Human CLN5 was originally described as a highly glycosylated, 60 kDa protein in BHK cells (Isosomppi, et al., 2002). More recently, a 52 kDa polypeptide was detected for overexpressed CLN5 in BHK and COS-7 cells by using a different antibody. Furthermore the study only reported high mannose-type sugars

for CLN5 (Lebrun, et al., 2009). In this study, using stably transfected CLN5-SH-SY5Y cells and the novel C-terminal antibody, we were able to detect the presence of both a 60 kDa and 50 kDa CLN5 polypeptides. Comprehensive analysis of maturation of CLN5 demonstrated that the 50 kDa polypeptide represents the mature form of the CLN5 protein, which is formed from the 60 kDa polypeptide most likely in the TGN or post-Golgi compartments. Although both forms were shown to contain complex-type sugars, only the 60 kDa polypeptide was affected by exposure to BFA, further supporting this hypothesis. The insensitivity of the 50 kDa form to BFA means it is absent from the ER/cis-Golgi compartments. The finding that the complete removal of N-glycans with PNGase F results in two polypeptides indicates that maturation of CLN5 involves additional modification(s). Among those, phosphorylation is one of the putative candidates, since the CLN5 amino acid sequence possesses many predicted sites for this modification.

Most soluble lysosomal proteins are targeted to the lysosomal compartment by mannose-6-phosphate-receptors (MPRs). It has been shown previously that CLN5 contains three putative Man-6-phosphorylation sites. Furthermore, using proteomics CLN5 has been shown to be secreted to the media of MPR deficient mouse fibroblasts (Kollmann, et al., 2005), supporting the MPR-dependent trafficking for CLN5. Here we demonstrate that mouse Cln5 could also be targeted to the lysosomes in the absence of MPRs (also human CLN5 (data not shown)). Our novel finding suggests that CLN5 can facilitate another route for its lysosomal trafficking in addition to the previously reported MPR-dependent pathway. A possible candidate for this alternative lysosomal trafficking could be the recently discovered LIMP-2 – mediated transport, that has been shown to be important for trafficking of glucocerebrosidase, a lysosomal enzyme defective in Gaucher disease (Reczek, et al., 2007). The multiligand receptor sortilin is also a good candidate for playing a role in the trafficking of CLN5, since it has been implicated in the targeting of another NCL protein, Cathepsin D (CLN10) (Canuel, et al., 2008). In addition to an MPR independent route, Cathepsin D can also use the classical MPR-pathway for its lysosomal trafficking (Canuel, et al., 2008; Kornfeld and Mellman, 1989; von Figura and Hasilik, 1986). Interestingly, palmitoyl protein thioesterase 1 (PPT1/CLN1), the protein defective in the infantile NCL, has also been reported to show properties that differ from those involved in the classical MPR-mediated sorting (Lyly, et al., 2007). Whether CLN5 can simultaneously utilize both MPR dependent and independent pathways to reach the lysosomes is currently unknown, as well as the extent of the contribution of each route to the trafficking of CLN5 under normal *in vivo* conditions. However, discovery and characterization of alternative transport pathways for CLN5, and other NCL proteins, in future experiments may provide important information to foster development of novel therapeutic strategies.

In this study, we also focused on the impact of several vLINCL disease-causing mutations in the trafficking and maturation of the CLN5 protein. Mutations in the *CLN5* gene have been reported to prevent trafficking of CLN5 to the level of ER (Lebrun, et al., 2009) and Golgi (Isosomppi, et al., 2002) but also not dramatically affect trafficking to the lysosomes (Vesa, et al., 2002). However, some of the previous data is conflicting (Isosomppi, et al., 2002; Vesa, et al., 2002), possibly as a result of the usage of different cell lines in the overexpression studies (COS, BHK, HeLa). Among the vLINCL patients, the existence of a genotype-phenotype correlation is not evident. However, p.R112H mutation has been reported in patients with a juvenile-onset (Pineda-Trujillo, et al., 2005), whereas the other four mutations (p.R112P, p.E253X, p.D279N, and p.Y392X) studied have been identified in patients exhibiting the classical clinical phenotype, the so-called Finnish vLINCL (Holmberg, et al., 2000). Also, it was recently reported that mutations in CLN5 (compound heterozygous p.C126Y and p.Y374C) result in late onset NCL (Sleat, et al., 2009). In this study, the mutations analyzed were found to result in ER localization of the mutated CLN5 at steady state, indicating that misfolding of the mutated proteins may be the major reason for the ER retention. However, none of the analyzed mutations were able to prevent the

proteolytical processing step of CLN5 in the ER, supporting the notion that the occurring cleavage is a signal sequence removal.

CLN5 carrying the classical onset causing mutation p.D279N was able to exit the ER and to some extent, even reach the lysosomes. Therefore, mutations that allow trafficking to the lysosomes possibly hit amino acids important for the function of CLN5 or disturb important interactions of the protein. Furthermore, CLN5 containing the later onset disease causing mutation p. R112P mainly localized to the ER and possibly to the Golgi complex, but was not found in the lysosomes. Therefore it is possible that CLN5 has a function already in the ER. In this regard, analysis of the glycosylation status of these mutated CLN5 polypeptides strengthens this hypothesis, as those retained in the ER also contained complex-type sugars, indicating previous trafficking along the Golgi complex. This would presuppose a retrograde transport of fully glycosylated CLN5 back to the ER, where it could interact with other ER residing proteins. That, in turn, may function as a prerequisite for the lysosomal transport of CLN5. Besides, retrograde transport would explain the fact that in the earlier studies wild-type CLN5 has been reported to partially reside in the ER (Holmberg, et al., 2004; Isosomppi, et al., 2002). As such the ER compartment is likely to represent an important place not only for CLN5 maturation/trafficking but maybe also for NCL interactions. These findings could also explain why localization of the mutated polypeptides does not always correlate to the severity of the disease phenotype.

The present study has confronted the existing discrepant information about the CLN5 protein and we have thoroughly analyzed the maturation and intracellular trafficking of this protein. Additionally, our novel finding of CLN5 utilizing an alternative pathway for lysosomal trafficking, implicates that it possesses characteristics different from classical lysosomal enzymes utilizing the MPR-mediated pathway. Naturally, this finding will affect the design of therapeutic strategies. Furthermore, we analyzed the consequences of disease mutations in the maturation and intracellular trafficking of CLN5. Although the primary function of CLN5 is still elusive, we could demonstrate that the disease mutations severely affect the lysosomal targeting of CLN5. The disease onset does not directly correlate to lysosomal localization of CLN5 and therefore, it is possible that CLN5 also functions in the early secretory pathway. The further description of normal and mutant CLN5 protein will pave way for further characterization of this protein, important for the development and functioning of the central nervous system.

ACKNOWLEDGEMENTS

The authors wish to express their gratitude to Auli Toivola for technical assistance and Carina von Schantz-Fant for critical discussion of this work. We thank Jussi Naukkarinen for revising the language. This study was financially supported by the Sigrid Juselius Foundation (to A.J.), the center of Excellence in Disease Genetics of the Academy of Finland Grant (213506 to A.J.), the Helsinki Biomedical Graduate School and Fundação para a Ciência e a Tecnologia (Portugal, grants FCT/POCI/SAU-MMO/55374/2004 and FCT/SFRH/BD/17560/2004 to C.B.).

FIGURES

Figure 1: Processing of CLN5. (A) The human CLN5 protein sequence possesses four possible initiation methionines (M1-4) at the amino acid positions of 1, 30, 50 and 62, and one predicted signal peptide cleavage site. N-terminal GFP-tag was added in front of all possible starting methionines M1-4. These constructs were then used to transiently transfect HeLa and COS-1 cells. COS-1 cell lysates were analyzed through SDS-PAGE and immunoblotting with the C-terminal CLN5 specific antibody (B) and with an anti-GFP antibody (C). Proteolytic cleavage of the GFP-(M1-M4) CLN5 proteins in COS-1 cells was verified by comparing the CLN5-specific staining pattern to that of the GFP by using a confocal microscope. LAMP-1 staining was used to indicate lysosomes (D, GFP.M2 shown as an example). Intracellular localization of the expressed GFP-constructs were also

analyzed in HeLa cells by staining with organelle-specific markers, PDI to indicate the ER (E, GFP.M2 shown as an example). Scale bar - 10 µm.

Figure 2: Maturation of human CLN5. SH-SY5Y cells, stably transfected with wild type CLN5 (CLN5-SH-SY5Y), were plated for incubation with either 50 µg/ml of cycloheximide (A) or 10 µg/ml of tunicamycin (C) for a period of 6 and 24 hours. Untransfected SH-SY5Y (ctrl) and untreated (ϕ) CLN5-SH-SY5Y cell lysates were used as controls. The pro- (p), mature (m) and glycosylated forms of CLN5 are indicated by arrow heads. The maturation was also studied in pulse-chase experiments using 100 µCi of [^{35}S] methionine (B). After a pulse of 1h, the cells were chased for the indicated time points and afterwards immunoprecipitated with CLN5-C/32 antibody. Protein glycosylation was also evaluated by digestion of CLN5-SH-SY5Y cell lysates with EndoH (D) or with peptide *N*-glycosidase F (PNGaseF) (E). Cells were treated with 2.5 µg/ml BFA for 5 h. Treatment of CLN5-SH-SY5Y cells with Brefeldin A leads to a downshift of the 60 kDa polypeptide (F). Additionally, cells were digested with EndoH or with PNGaseF and immunoblotted. In all shown Western Blots (A-E) CLN5 was detected with a novel C-terminal human CLN5 antibody (CLN5-C/32) described in Materials and Methods. The CLN5-C/32 antibody also recognized an unspecific band in Western Blot, which is indicated with an asterisk (*).

Figure 3: CLN5 is able to traffic to lysosomes in the absence of Mannose-6-phosphate receptors. Mouse MPR deficient cells were transiently co-transfected with myc-tagged mouse Cln5 (see Materials and Methods) and human CLN3 constructs. The cells were immunostained with CLN5 (A) and CLN3 (B) specific antibodies and as a lysosomal marker, with a LAMP-1 antibody (C). Cells were analyzed with a confocal immunofluorescence microscope. Co-localization is indicated by: CLN5 and LAMP-1 – pink; CLN5 and CLN3 - yellow; CLN5, CLN3 and LAMP-1 – white (D and insert). Scale bar - 10 µm.

Figure 4: Steady state localization of WT and mutated CLN5 in HeLa cells. A scheme illustrates the relative locations and amino acid changes of the analyzed mutations in CLN5 (GenBank NM_006493.2) that result in vLINCL. Wt CLN5 (A-C) and CLN5 constructs carrying the mutation p. R112P (POR) (D-F), p. R112H (COL), which results in a later onset vLINCL (G-I) or p. E253X (SWE) (J-L) were transiently transfected into HeLa cells. Localization of the different CLN5 proteins was studied by immunofluorescence labelling, followed by confocal microscopy using a CLN5 specific antibody (B, E, H, K). PDI was used to label the ER (D, G, J) and LAMP-1 was used to detect lysosomes (A). Co-localization is indicated in yellow. Scale bar - 10 µm.

Figure 5: Localization of the mutants after chx treatment in HeLa cells

CLN5 constructs carrying the mutations p. R112P (POR) (A-C), p. R112H (COL) (D-F), or p. D279N (EUR) (G-I), p. E253X (SWE) (J-L) or p. Y392X (Fin_{Major}) (M-O) were transiently transfected into HeLa cells. 48 h posttransfection the cells were treated for 1-5 h with cycloheximide (chx) and afterwards fixed with ice cold methanol. Localization of the different CLN5 proteins was studied by immunofluorescence labelling, followed by confocal microscopy using a CLN5 specific antibody (B, E, H, K, N). PDI was used to label the ER (A, D, G) and LAMP-1 was used to detect lysosomes (J, M). Co-localization is indicated in yellow. Scale bar - 10 µm.

Figure 6: Maturation of the CLN5 polypeptides carrying vLINCL disease causing missense mutations

COS-1 cells were transiently transfected with CLN5 constructs carrying the missense mutations p. R112H (COL), p. R112P (POR) or p. D279N (EUR) for 48 h. Cell lysates were digested with EndoH to remove high-mannose-type sugars or with PNGaseF to remove almost all types of N-glycans including complex sugars. Untreated (ϕ) and treated samples were then immunoblotted and CLN5 was detected with a novel C-terminal human CLN5 antibody

(CLN5-C/32) described in the Materials and Methods. Unspecific bands are indicated with an asterisk (*).

REFERENCES

Bendtsen JD, Nielsen H, von Heijne G, Brunak S. 2004. Improved prediction of signal peptides: SignalP 3.0. J Mol Biol 340(4):783-95.

Bessa C, Teixeira CA, Mangas M, Dias A, Sa Miranda MC, Guimaraes A, Ferreira JC, Canas N, Cabral P, Ribeiro MG. 2006. Two novel CLN5 mutations in a Portuguese patient with vLINCL: insights into molecular mechanisms of CLN5 deficiency. Mol Genet Metab 89(3):245-53.

Canuel M, Korkidakis A, Konnyu K, Morales CR. 2008. Sortilin mediates the lysosomal targeting of cathepsins D and H. Biochem Biophys Res Commun 373(2):292-7.

Haltia M. 2006. The neuronal ceroid-lipofuscinoses: from past to present. Biochim Biophys Acta 1762(10):850-6.

Holmberg V, Jalanko A, Isosomppi J, Fabritius AL, Peltonen L, Kopra O. 2004. The mouse ortholog of the neuronal ceroid lipofuscinosis CLN5 gene encodes a soluble lysosomal glycoprotein expressed in the developing brain. Neurobiol Dis 16(1):29-40.

Holmberg V, Lauronen L, Autti T, Santavuori P, Savukoski M, Uvebrant P, Hofman I, Peltonen L, Jarvela I. 2000. Phenotype-genotype correlation in eight patients with Finnish variant late infantile NCL (CLN5). Neurology 55(4):579-81.

Isosomppi J, Vesa J, Jalanko A, Peltonen L. 2002. Lysosomal localization of the neuronal ceroid lipofuscinosis CLN5 protein. Hum Mol Genet 11(8):885-91.

Jalanko A, Braulke T. 2008. Neuronal ceroid lipofuscinoses. Biochim Biophys Acta.

Jarvela I, Sainio M, Rantamaki T, Olkkonen VM, Carpen O, Peltonen L, Jalanko A. 1998. Biosynthesis and intracellular targeting of the CLN3 protein defective in Batten disease. Hum Mol Genet 7(1):85-90.

Kollmann K, Mutenda KE, Balleininger M, Eckermann E, von Figura K, Schmidt B, Lubke T. 2005. Identification of novel lysosomal matrix proteins by proteome analysis. Proteomics 5(15):3966-78.

Kornfeld S, Mellman I. 1989. The biogenesis of lysosomes. Annu Rev Cell Biol 5:483-525.

Kyttala A, Lahtinen U, Braulke T, Hofmann SL. 2006. Functional biology of the neuronal ceroid lipofuscinoses (NCL) proteins. Biochim Biophys Acta 1762(10):920-33.

Lebrun AH, Storch S, Ruschendorf F, Schmiedt ML, Kyttala A, Mole SE, Kitzmuller C, Saar K, Mewasingh LD, Boda V and others. 2009. Retention of lysosomal protein CLN5 in the endoplasmic reticulum causes neuronal ceroid lipofuscinosis in Asian Sibship. Hum Mutat.

Luiro K, Kopra O, Lehtovirta M, Jalanko A. 2001. CLN3 protein is targeted to neuronal synapses but excluded from synaptic vesicles: new clues to Batten disease. Hum Mol Genet 10(19):2123-31.

Lyly A, von Schantz C, Salonen T, Kopra O, Saarela J, Jauhiainen M, Kyttala A, Jalanko A. 2007. Glycosylation, transport, and complex formation of palmitoyl protein thioesterase 1 (PPT1)--distinct characteristics in neurons. BMC Cell Biol 8:22.

Pineda-Trujillo N, Cornejo W, Carrizosa J, Wheeler RB, Munera S, Valencia A, Agudelo-Arango J, Cogollo A, Anderson G, Bedoya G and others. 2005. A CLN5 mutation causing an atypical neuronal ceroid lipofuscinosis of juvenile onset. Neurology 64(4):740-2.

Pohlmann R, Boeker MW, von Figura K. 1995. The two mannose 6-phosphate receptors transport distinct complements of lysosomal proteins. J Biol Chem 270(45):27311-8.

Reczek D, Schwake M, Schroder J, Hughes H, Blanz J, Jin X, Brondyk W, Van Patten S, Edmunds T, Saftig P. 2007. LIMP-2 is a receptor for lysosomal mannose-6-phosphate-independent targeting of beta-glucocerebrosidase. Cell 131(4):770-83.

Santavuori P. 1988. Neuronal ceroid-lipofuscinoses in childhood. Brain Dev 10(2):80-3.

Savukoski M, Klockars T, Holmberg V, Santavuori P, Lander ES, Peltonen L. 1998. CLN5, a novel gene encoding a putative transmembrane protein mutated in Finnish variant late infantile neuronal ceroid lipofuscinosis. Nat Genet 19(3):286-8.

Sleat DE, Ding L, Wang S, Zhao C, Wang Y, Xin W, Zheng H, Moore DF, Sims KB, Lobel P. 2009. Mass spectrometry-based protein profiling to determine the cause of lysosomal storage diseases of unknown etiology. Mol Cell Proteomics 8(7):1708-18.

Sleat DE, Zheng H, Qian M, Lobel P. 2006. Identification of sites of mannose 6-phosphorylation on lysosomal proteins. Mol Cell Proteomics 5(4):686-701.

Vesa J, Chin MH, Oelgeschlager K, Isosomppi J, DellAngelica EC, Jalanko A, Peltonen L. 2002. Neuronal ceroid lipofuscinoses are connected at molecular level: interaction of CLN5 protein with CLN2 and CLN3. Mol Biol Cell 13(7):2410-20.

von Figura K, Hasilik A. 1986. Lysosomal enzymes and their receptors. Annu Rev Biochem 55:167-93.

von Heijne G. 1986. A new method for predicting signal sequence cleavage sites. Nucleic Acids Res 14(11):4683-90.

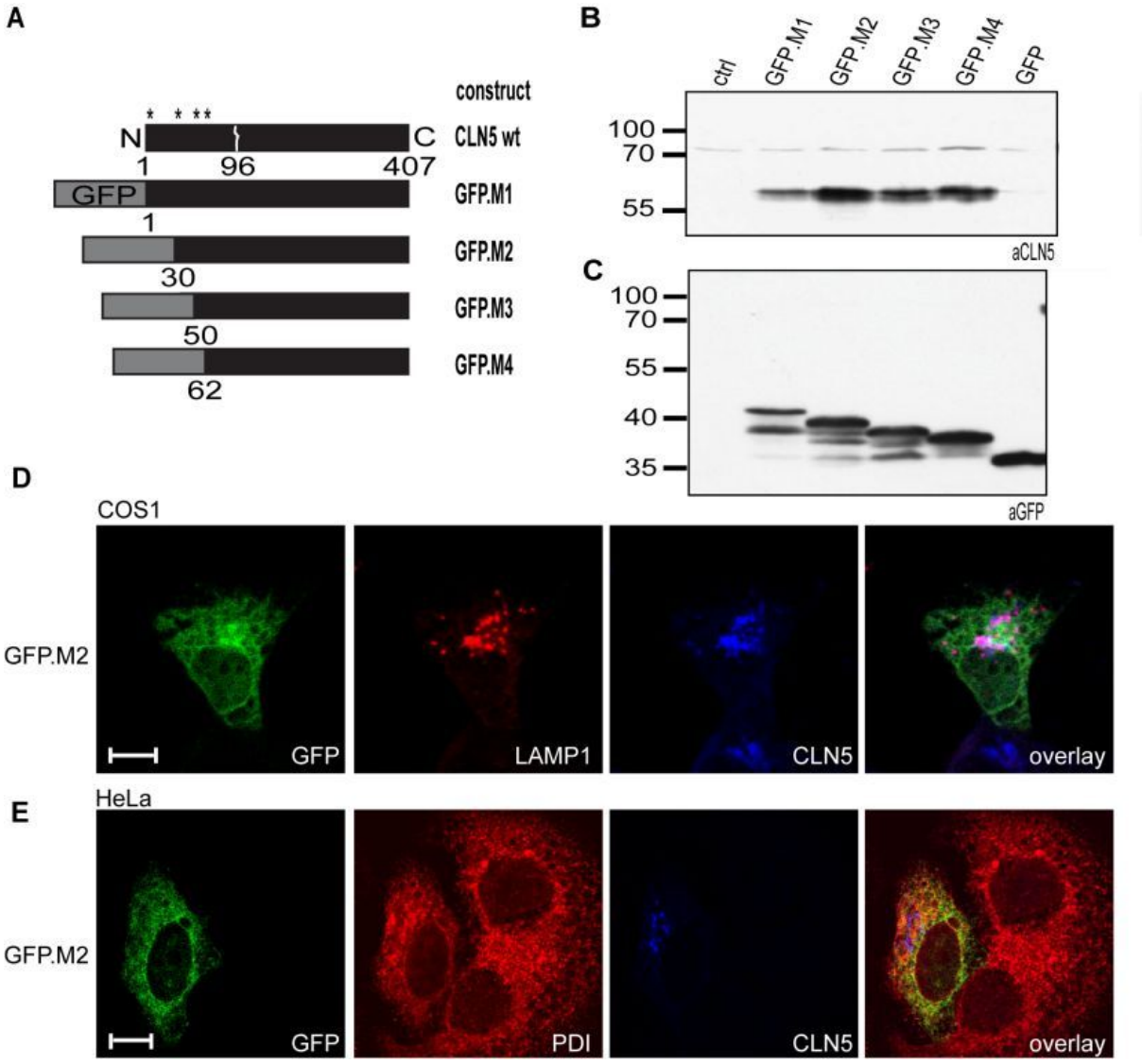

A
construct
N
C
CLN5 wt
1
96
407
GFP
GFP.M1
1
GFP.M2
30
GFP.M3
50
GFP.M4
62
B
ctrl
GFP.M1
GFP.M2
GFP.M3
GFP.M4
GFP
100
70
55
aCLN5
C
100
70
55
40
35
aGFP
D
COS1
GFP.M2
GFP
LAMP1
CLN5
overlay
E
HeLa
GFP.M2
GFP
PDI
CLN5
overlay

Figure 1.

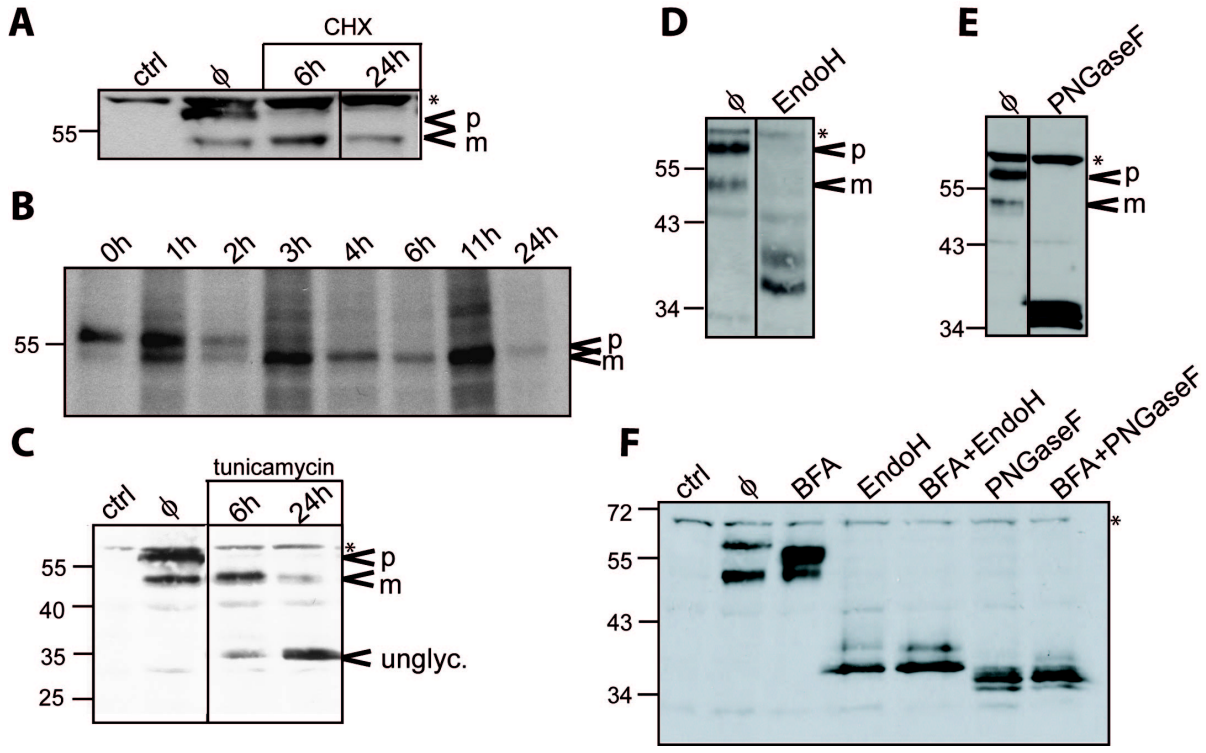


Figure 2

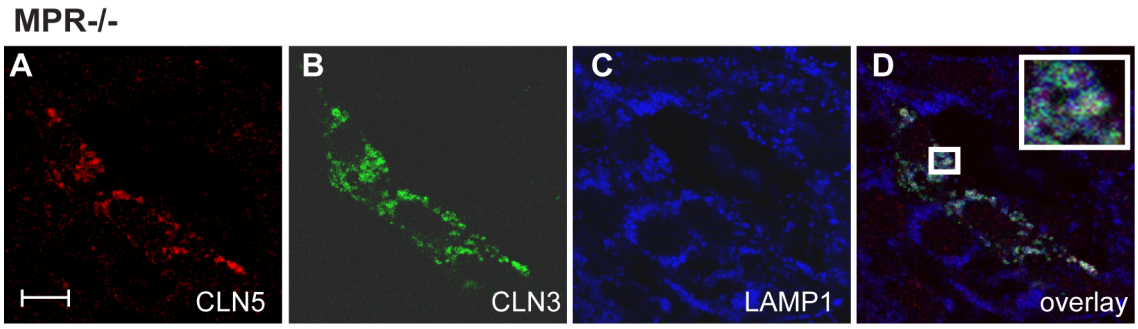


Figure 3

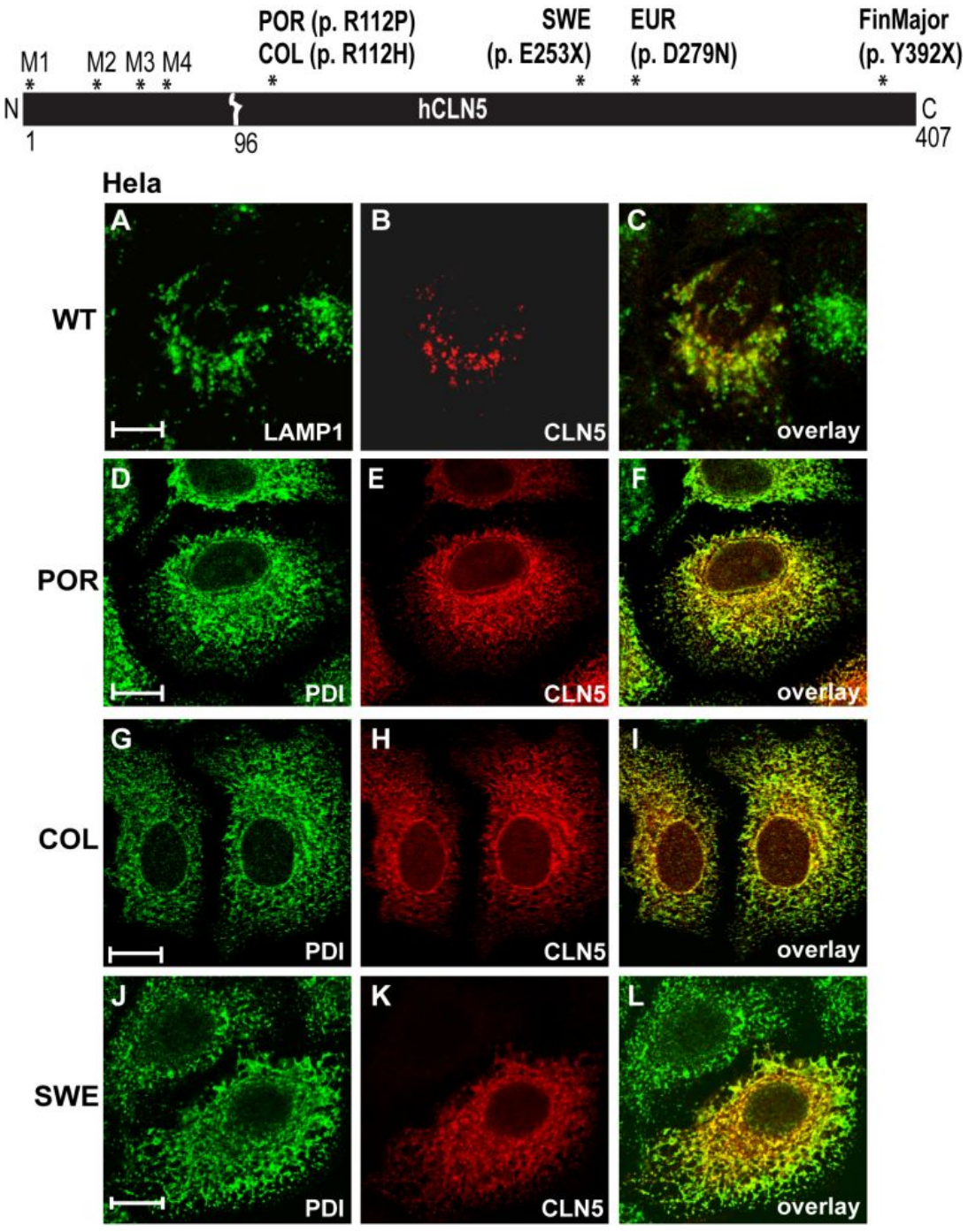

POR (p. R112P)
COL (p. R112H)
SWE
(p. E253X)
EUR
(p. D279N)
FinMajor
(p. Y392X)
M1
M2 M3 M4
N
hCLN5
C
1
96
407
Hela
WT
POR
COL
SWE
A
B
C
D
E
F
G
H
I
J
K
L
LAMP1
CLN5
overlay
PDI
CLN5
overlay
PDI
CLN5
overlay
PDI
CLN5
overlay

Figure 4

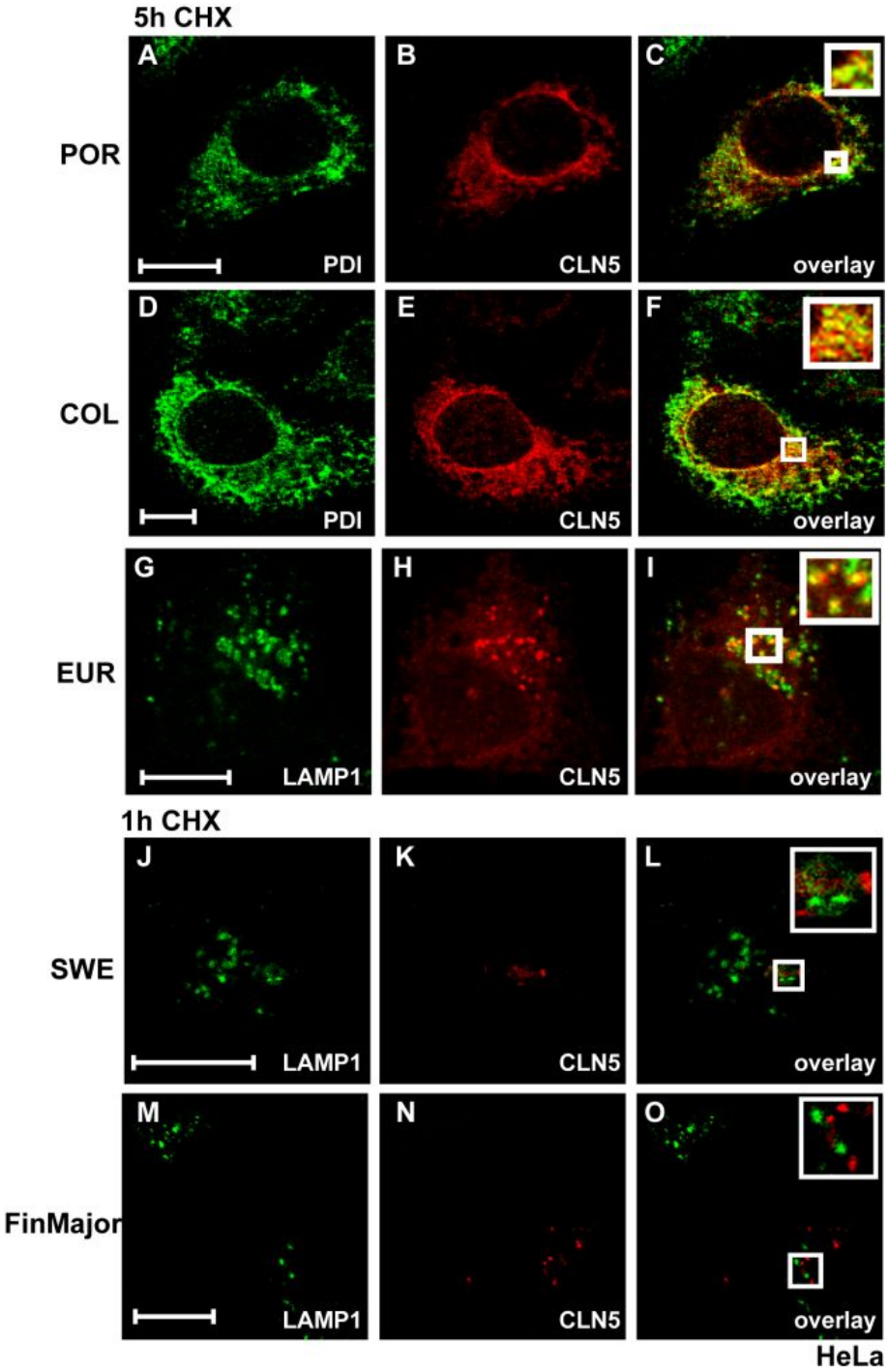

5h CHX
A
B
C
POR
PDI
CLN5
overlay
D
E
F
COL
PDI
CLN5
overlay
G
H
I
EUR
LAMP1
CLN5
overlay
1h CHX
J
K
L
SWE
LAMP1
CLN5
overlay
M
N
O
FinMajor
LAMP1
CLN5
overlay
HeLa

Figure 5

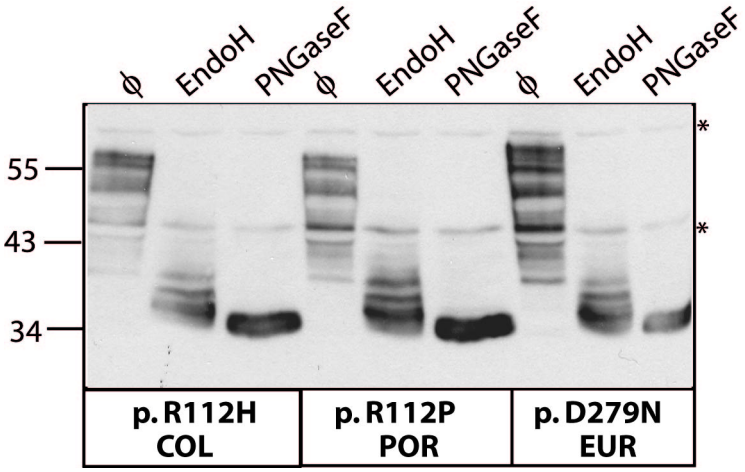

ϕ
EndoH
PNGaseF
ϕ
EndoH
PNGaseF
ϕ
EndoH
PNGaseF
55
43
34
*
*
p. R112H
COL
p. R112P
POR
p. D279N
EUR

Figure 6

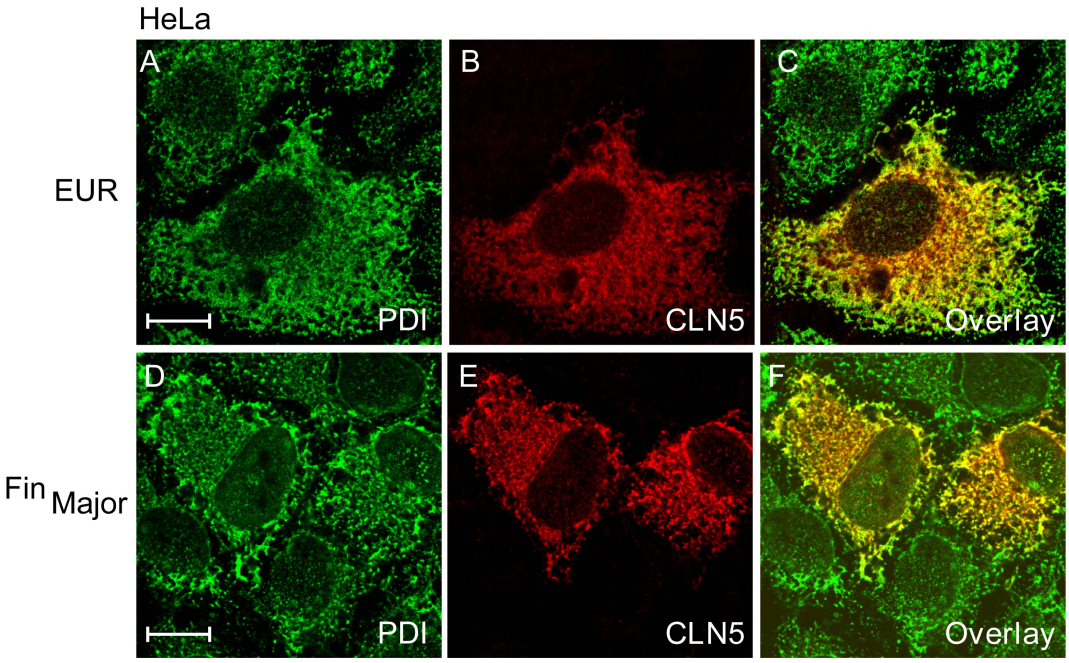
SUPPORTING INFORMATION

program		hCLN5 M1	hCLN5 M2	hCLN5 M3	hCLN5 M4	mCln5
SignalP	signal peptide	No	Yes	Yes	Yes	Yes
	probability	0,143	0,785	0,987	1	0,999
	aa position (pos.)	-	62 or 66	46	34	33
	aa pos. (M1)	-	92 or 96	96	96	-
SigCleave	score	9,11/7,55	9,11/7,55	9,11/7,55	9,4/7,5	8,78/7,75
	aa pos.	89/91	60/62	40/42	28/30	27/28
	aa pos. (M1)	89/91	90/92	90/92	90/92	-

Supp. Table S1: Signal peptide prediction. The prediction of signal peptides and the correspondent cleavage sites in human CLN5 sequences (M1-M4) and mCln5 are shown in this table. The position of the predicted signal peptidase I cleavage site is represented in reference to both the initiation methionine used in the analyzed sequence but also to the sequence starting with M1. Additionally the probability of a signal peptide and the overall verdict is also presented for the SignalP model. For the SigCleave prediction only the two highest scores are presented.

antigen	name	produced in	detected localization in IF	approximate molecular weight in WB	reference
CLN5 aa 1-75	1RI-3	rabbit	ER	50 kDa	Vesa et al., 2002
CLN5 aa 105-119	33	rabbit	lysosomes (our unpublished data)	49 kDa	Bessa et al., 2006
CLN5 aa 258-273	5289	rabbit	lysosomes	60 kDa	Isosomppi et al., 2002
CLN5 aa 393-407	CLN5-C/32	rabbit	lysosomes	50 kDa + 60 kDa	Described here
GST-Cln5 aa 40-341	1RmII-4	rabbit	lysosomes	-	Holmberg et al., 2004
GST-Cln5 aa 40-341	1GmII-3 2GmII-3	guinea pig	lysosomes	50 kDa + 60 kDa (our unpublished data)	Holmberg et al., 2004; Lebrun et al., 2009; used here

Supp. Table S2: Overview of CLN5 antibodies. The antibodies used for the recognition of CLN5 are listed in a table. Indicated are also the detected intracellular localization of CLN5 protein in immunfluorescence and the reported molecular weights detected in Western Blot studies.



Supp. Figure S3: Steady state localization of the previously characterized CLN5 mutations in HeLa cells. CLN5 constructs carrying the mutation p. D279N (EUR) and p. Y395X (Fin_{Major}) were transiently transfected into HeLa cells. Localization of the different CLN5 proteins was studied by immunofluorescence labelling, followed by confocal microscopy using a CLN5 specific antibody (B, E). PDI was used to label the ER (A, D). Co-localization is indicated in yellow. Scale bar - 10 µm.