ROLE OF N-GLYCOSYLATION IN TRAFFICKING AND STABILITY OF HUMAN CLN5

By

AKSHAY MOHARIR

B.S., Rashtrasant Tukdoji Maharaj Nagpur University, 2005

M.S., Cochin University of Science and Technology, 2007

A THESIS

Submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Division of Biology
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2012

Approved by:

Major Professor

Dr. Stella Y lee

ABSTRACT

Neuronal Ceroid Lipofuscinoses (NCLs) are a group of lysosomal storage diseases that are characterized by accumulating autofluorescent lipopigments in cells. NCLs are a form of progressive neurodegenerative diseases with symptoms ranging from blindness, loss of speech and motor activities to ataxia and seizures. Patients do not live to adulthood in most cases, making it prevalent in children. Among the many genes that cause NCL, *CLN5* leads to different forms of NCL (infantile, late infantile, juvenile, and adult). CLN5 protein resides in the lysosomes but its function has not been established. It is predicted to contain eight N-glycosylation sites, but the role of N-glycosylation on its function and trafficking has not been assessed.

We analyzed the role of N-glycosylation on the transport and stability of human CLN5. We created N-glycosylation mutants of each site by changing the Asn to Gln and our analysis of these mutants show that all the eight N-glycosylation sites are used *in vivo*. We also report effects of abolishing individual N-glycosylation sites on the trafficking of CLN5. While the lack of glycosylation at some sites results in CLN5 being retained in the ER or Golgi, others do not affect CLN5 trafficking. Cycloheximide chase experiments show that one of the mutants (N401Q) in CLN5 leads to lower protein levels in cell pellets with an increased secretion compared to CLN5 wild type, while other mutations show differential stability in cell pellets. These results demonstrate that each N-glycosylation site plays a different role(s) in the stability, transport and/or function of CLN5.

Table of Contents

List of figures iv

List of tables v

Chapter 1 – Literature review 1

Biosynthetic transport pathway.. 2

Cargo proteins 5

Transport machinery 6

Lysosomal acid hydrolase transport 9

Lysosomal storage diseases 10

Neuronal Ceroid Lipofuscinosis 12

CLN5 16

References 18

Chapter 2 – Role of N-glycosylation in trafficking and stability of human CLN5 27

Abstract 28

Introduction 29

Materials and Methods 32

Results 38

N-glycosylation of CLN5 *in vivo* 38

Subcellular localization of CLN5 N-glycosylation mutants 39

Deglycosylation studies of CLN5 41

Stability of CLN5 deficient in N-glycosylation 43

Subcellular localization of CLN5 expressed in patients 44

Discussion 44

Figures and Tables 49

References 75

Chapter 3 - Discussion and Future prospective 77

List of Figures

Chapter 1 – Literature review

Figure 1.1 – General view of biosynthetic transport pathway …………………. 3

Chapter 2 – Role of N-glycosylation in trafficking and stability of human CLN5

Figure 2.1 – N-glycosylation of CLN5 *in vivo* ……………………………………………. 49

Figure 2.2 – Subcellular localization of CLN5 N-glycosylation mutants …..... 51

Figure 2.3 – Accumulation of lysosomal localized CLN5 in rab5A Q79L- induced enlarged endosomes ……………………………………………………………………………… 56

Figure 2.4 – Endo H sensitivity and deglycosylation studies of CLN5 ……….. 63

Figure 2.5 – Stability of N-glycosylation deficient CLN5 protein ………………… 66

Figure 2.6 – Subcellular localization of CLN5 patient mutations ………………… 68

Supplementary figure S1 – Subcellular localization of CLN5 wt …………………. 70

Supplementary figure S2 – Sequence alignment of CLN5 ………………………….. 72

List of Tables

Chapter 1 – Literature review

Table 1.1 – List of NCL genes, symptoms and defects …………………………………… 14

Chapter 2 – Role of N-glycosylation in trafficking and stability of human CLN5

Table 2.1 – List of primers sequences used ………………………………………………….. 37

Table 2.2 – Summary of subcellular localization of CLN5 N-glycosylation mutants 61

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my major professor, Dr. Stella Y. Lee for giving me an opportunity to work on this project. I would also like to express my appreciation to my committee members: Dr. Jeroen Roelofs and Dr. Lorena Passarelli for their constructive criticism and advice on the project.

I would like to thank Dr. Roman Ganta, Dr. Subbarath Muthukrishnan and Dr. Brian Spooner for their invaluable support and advice during tough times. I am grateful for the wonderful friends I have and thank them for their support and for making my stay in Manhattan feel like home.

Finally I would like to thank my family for their unending love and encouragement, without which none of this would have been possible.

This project has been supported by start-up fund-Biology (KSU), K-INBRE recruitment package and Johnson Center for Basic Cancer Research.

CHAPTER 1

LITERATURE REVIEW

BIOSYNTHETIC TRANSPORT PATHWAY

Proteins play an important role in organelle dynamics and lend functional and structural characteristic features to the organelles they reside in. These proteins serve specific functions in these organelles and the loss of correct localization of these proteins can cause diseases. The process of delivering the proteins to their correct destination starts with their synthesis in the Endoplasmic Reticulum (ER). A newly synthesized protein has to be delivered to its specific cellular location which is accomplished by its transport through a series of membrane bound compartments. This transport process forms an essential part of intracellular protein trafficking which involves various adaptor proteins that aid in transport of cargo proteins to their destination (1). Mitochondrial proteins are an exception and do not follow this path. Instead, they are delivered to the mitochondria directly from the cytoplasm via other mechanisms/pathways. There are two basic pathways that have been broadly classified for protein trafficking: the biosynthetic pathway and the endocytic pathway. The biosynthetic or the secretory pathway is involved in the transport of proteins from the ER to the Golgi and eventually to their final destination (lysosome, plasma membrane or secreted out of cell). In general, the process of moving proteins outward is termed anterograde transport (2, 3). This movement along the biosynthetic pathway is not random and involves various transporters and signal sequences that may be present in or acquired by the protein. Some proteins (such as lysosomal hydrolases) acquire the sorting signal through post translational modification in the Golgi. A generalized view of the biosynthetic pathway is shown in figure 1.1. The endocytic pathway on the other hand involves the movement of proteins from the plasma membrane to

Figure 1.1

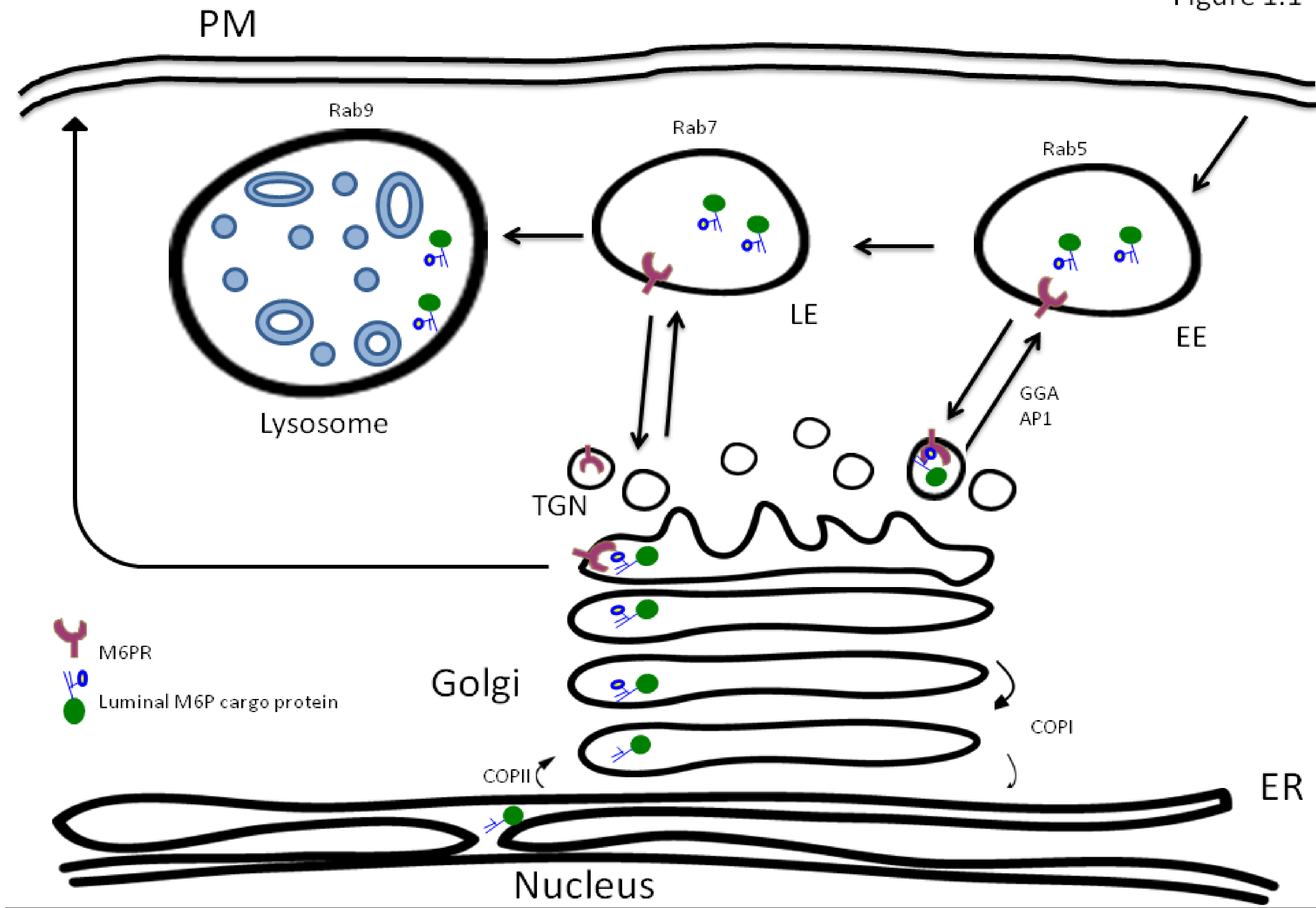


Figure1.1- General view of the biosynthetic pathway. The biosynthetic pathway is shown in a simplified manner. A newly synthesized cargo protein destined for the lysosome lumen is trafficked from the ER to the Golgi, where it acquires the M6P tag in the cis Golgi which helps in recognition by the M6PR in TGN. The cargo protein in then transported to the endosomes by the M6PR where the M6PR dissociates from the cargo protein and recycles back to the TGN. The endosome matures to form the lysosome with the cargo protein in the lumen. PM – Plasma Membrane, LE- Late Endosome, EE- Early Endosome, TGN – Trans Golgi network, ER- Endoplasmic Reticulum, Rab5 – EE marker, Rab7- LE marker, Rab9-Lysosome marker, COPI – Coatomer protein complex-I, and COPII – Coatomer protein complex-II.

endosomes and then to the lysosomes for degradation. The process of moving proteins inward is generally termed retrograde transport. The endocytosed proteins become a part of the endocytic pathway. (4)

Cargo Proteins

The proteins being transported to their final destination through different pathways are termed cargo proteins. These cargo proteins can be soluble luminal proteins (e.g., lysosomal hydrolases) or transmembrane proteins (e.g., signaling receptors). Depending on the type of protein, different signals and adaptor/transport molecules aid in the proper targeting through organelles to their final destination. Usually soluble luminal proteins are recognized and bound by receptors that propel their transport (e.g., Mannose 6-phosphate receptors or M6PR) while membrane proteins contain signaling sequences that serve the same purpose. The luminal and transmembrane cargo proteins undergo various post translational modifications (in the ER, Golgi and/or lysosomes) that can serve different purposes. The most common modifications include glycosylation which are of two prevalent types. Glycosylation on Asn is termed N-linked glycosylation or N-glycosylation and starts in the ER. Another type of glycosylation is on Ser/Thr residue (termed O-linked glycosylation) which normally occurs in the Golgi. In addition, glycosylated proteins undergo modifications on their N- or O-linked glycans (e.g., phosphorylation on mannose, and sialylation) (5). These modifications serve various functions including proper recognition of the cargo proteins by their transporters. In addition, some modifications (N- and O-Glycosylation) protect the cargo protein from being cleaved or degraded (e.g., lysosomal luminal proteins are protected from lysosomal hydrolases). These

modifications can either be removed once the cargo protein reaches its destination (e.g., Mannose 6-Phosphate) or remain unchanged (e.g., N-glycosylation).

Transport Machinery

The exit of luminal or transmembrane proteins from the ER involves their concentration into vesicles that are coated with Coatomer protein complex II (COPII) (6). Transmembrane cargo proteins exiting the ER contain an ER exit sequence which enables an organized exit (7, 8). These signals can be either di-acidic (YTDIE), di-basic (RR) or di-hydrophobic (FF and LV). The components of COPII recognize these motifs and bind them, enabling their transport out of the ER and to the Golgi (4, 9). COPII recruitment to the ER exit site is propelled by the activation of SAR1 (a small GTPase) (4). Some proteins (soluble luminal proteins) that need to be retained in the ER lumen are also transported to the Golgi as part of the bulk flow, but they contain an ER retrieval signal at the C- terminal (KDEL) which aids in their retrieval from the Golgi to the ER (10). Type I membrane proteins of the ER contain a different type of retrieval signal (KKXX) that enables them to be retrieved to the ER (11). The ER-Golgi intermediate complex (ERGIC) is an essential location for retrieval of ER resident proteins (12) . This retrograde transport of proteins is mediated by COPI coated vesicles which are recruited similar to COPII by another small GTPase called ARF1 (12).

Once the cargo proteins leave the ER, they move to the Golgi. The Golgi has a stack like structure consisting of cis, medial and trans compartments/cisternae with the cis towards the ER and trans facing the plasma membrane (PM). The cargo within the Golgi is thought to progress through the different stacks and eventually bud off from the trans Golgi. The model

currently accepted which explains the maturation of Golgi stacks is the cisternal maturation model. This model proposes that the budding vesicles from ER fuse to form the new cis cisternae while the old cis cisternae matures into the medial cisternae and the medial cisternae to the trans cisternae. The Trans Golgi network (TGN) acts as a hub for various post translational modifications and for sorting proteins for their final destination (5). The transport of cargo proteins from the Golgi to endosomes or PM occurs via small vesicles that bud from the TGN (donor) and merge with acceptor compartments. The most common coat protein that aids in transport via vesicles is clathrin. The vesicles that bud from the Golgi are coated with clathrin and it is interesting that only the last cisternae in the TGN have a clathrin coat while the earlier cisternae have different coat proteins (13). This might be an indication of differential trafficking phenomena in different cisternae of the TGN. Since transport to the endosomes is mediated by clathrin coated vesicles/ tubules, the last cisternae of the TGN might be involved in this process while other earlier cisternae may function in sorting proteins to other compartments or destinations.

Exiting the Golgi or the TGN also involves specific signals and adaptors similar to or more complex than ER exit. The next destination after TGN in the biosynthetic pathway is usually the PM or endosome (14). The best examples of cargo proteins sorted from the TGN to the endosomes based on their signals are lysosomal hydrolases that have mannose 6-phosphate (M6P) residues on them. These M6P residues are recognized by either cationic-dependent (CD-M6PR) or -independent (CI-M6PR) mannose 6- phosphate receptors to target the cargo protein to the endosomes. The cargo receptors in the TGN can utilize two types of signals for transporting cargo to the endosomes. One type of signal is present in the cytoplasmic domain

of the cargo receptor and consists of sequences that bind to AP1 (an adaptor complex of the TGN) (15). The other type of signal that is utilized by proteins such as β-secretase and sortilin is DXXLL. This signal is recognized by GGA (Golgi localized, γ-ear containing, ADP ribosylation factor (ARF)-binding protein) adaptors (14). AP1 and GGA adaptors aid in vesicle formation and transport to the endosomes.

The endosomes are an integral part of the endocytic pathway as well as the biosynthetic pathway. Different forms of endosomes have been observed that undergo maturation and eventually fuse with the lysosomes. Lysosomal proteins (luminal and membrane) traffic from the TGN to the lysosome via endosomes (16). Endosomes can be early endosomes (EE) or late endosomes (LE). The late endosomes fuse with lysosomes to form intermediates called endolysosomes that eventually mature into lysosomes. The EE and LE can be distinguished by different markers, mainly belonging to the Ras superfamily of Rab GTPases. The Rab5 is characteristic marker of EE which switches to Rab7 in the LE and eventually to Rab9 in the lysosomes (17, 18).

Lysosomal membrane cargo proteins have sorting motifs that are highly conserved and adaptor proteins recognize these motifs and bind to them. An example is the family of tyrosine based motifs YXXϕ (X is any amino acid (aa) and ϕ is an aa with a bulky hydrophobic side chain) (19). These motifs have been shown to interact with all known adaptor complexes by yeast two-hybrid studies (20-23). The YXXϕ motifs are known to be involved in the sorting to endosomes and lysosomes from the TGN and from PM to the TGN (19). Another family of motifs which is involved in similar functions as YXXϕ is the dileucine based motifs. These motifs

are recognized by adaptor proteins AP-1, AP-2 and AP-3 (24). These dileucine motifs are surrounded by acidic residues which are critical for their functioning (25).

Lysosomal acid hydrolase transport

The best studied examples of lysosomal/endosomal protein transport are lysosomal acid hydrolases. Soluble lysosomal proteins have their signal peptide cleaved as they enter the ER after which they undergo N-glycosylation where a preformed oligosaccharide core (Glc3Man9GlcNac2) is added onto the asparagine residue within the consensus sequence N-X-S/T (X is any aa residue except Pro or Asp). This modification can be very important in protecting the hydrolases from being cleaved in the acidic environment, in addition to stabilizing the protein as well as help in sorting and transport (26). Once these soluble lysosomal proteins leave the ER and arrive in the Golgi, further modifications take place on the oligosaccharide that was added in the ER. The sugars are either trimmed or further addition of complex sugars can take place (e.g., fucose or N-acetyl neuraminic acid) or there is phosphorylation of select mannose residues (M6P). The M6P tag is added by the action of two enzymes – GlcNac-1-phosphotransferase and uncovering enzyme (UCE) (27-29). M6P receptors (M6PR) bind to the M6P in the TGN and transport the proteins to the endosomes where the acidic pH causes them to dissociate from their cargo. The hydrolases are then transferred onto late endosomes and lysosomes. The M6PRs recycle back to the TGN to bind fresh cargo proteins (30, 31). As mentioned before, there are two types of M6PRs – CI-M6PR and CD-M6PR. It has been shown that in mice lacking CD-M6PR or CI-M6PR, the cells utilize other mechanisms to target the proteins properly to the lysosomes (32). There have been several alternative

receptors that have been reported to carry out similar functions as the M6PR. Sortilin has been suggested to be an alternative as it is involved in the transport of acid sphingomylinase and prosaposin to the lysosomes (33, 34). Another receptor, lysosomal integral membrane protein type 2 (LIMP-2) is involved in the lysosomal targeting of β-glucocerebrosidase (35). Sorting and trafficking of lysosomal hydrolases from the TGN is mediated by clathrin/ AP1 coated vesicle formation (25) and ARF1 is responsible for the recruitment of AP1 to the TGN membrane (36, 37). AP1 also plays a role in the process of retrograde transport from endosomes to TGN (38).

LYSOSOMAL STORAGE DISEASES (LSD)

Lysosomal storage diseases (LSD) are caused by defects in lysosomal protein function(s) or their trafficking, resulting in the accumulation or buildup of different kinds of material within the lysosomes. Lysosomes contain almost 50-60 soluble hydrolases (39) in addition to integral membrane proteins (40). An example of LSD, Gauchers disease is caused by defects in the transport of β-glucocerebrosidase (β-Glc). The transport of β-Glc to the lysosomes is dependent on LIMP2 (35) and mutations or deletion of LIMP2 leads to reduced activity of β-Glc (41) as its transport to the lysosome is impaired. Mutating proteins involved in the multivesicular body (MVB) pathway also causes trafficking problems as has been shown in a recent study in yeast which leads to change in the vacuolar pH (42). The maintenance of low pH is essential for the functioning of lysosomes because it aids in various processes such as loading of neurotransmitters onto vesicles and lysosomal hydrolase maturation (43, 44). Apart from that, the low pH is necessary for the release of protein cargo from the M6PRs and increasing

the pH can lead to the accumulation of M6PRs in the endosomal compartments and not recycling to the TGN (45).

Among the best studied LSD caused by defective transport of proteins to lysosomes is the I-cell disease or mucolipidosis II (ML II). The disease is caused due to a faulty M6PR system which results in defective transport of lysosomal enzymes. The defect is in the gene encoding the enzyme responsible for catalyzing the first step of formation of M6P (GlcNac-1-phosphotranferase), causing the secretion of lysosomal hydrolases to the outside of the cell as these hydrolases do not get properly modified to be recognized by the M6PR transport machinery. Another such disease is the Niemann-Pick C (NPC) which results due defective cholesterol transporter (NPC1) activity (40, 46).

There can be various proteins and lipids that can accumulate within the lysosomes in LSD, one of which is subunit c of mitochondrial ATP synthase which is seen in several LSDs (including Neuronal ceroid lipofuscinosis or NCL). Yeast cells that are deficient in btn1p (orthologue of CLN3) show altered vacuolar pH due to afflicted ATP synthase trafficking (47). It has been observed that usually LSD result in deficiency of enzyme(s) involved in the disease and cell compensates by upregulating the expression and synthesis of other lysosomal proteins. This is indicative of interdependence between the different lysosomal proteins (48). Studies elucidating the motif involved in the upregulation of lysosomal proteins during LSD point towards a ten base pair sequence close to the transcription start site called coordinated lysosomal expression and regulation (CLEAR). This element has been found in various lysosomal proteins such as CLN5, CLN3, LAMP-1, NPC1, NPC2 and β-galactosidase. Transcription factor EB (TFEB) binds to the CLEAR element and induces its transcription (48).

NEURONAL CEROID LIPOFUSCINOSIS (NCL)

NCLs are a group of clinically and genetically heterogeneous progressive neurodegenerative disorders that mostly affect children. They are usually inherited in an autosomal recessive manner (49). The age of onset of NCLs is variable with varying symptoms such as epilepsy, mental and motor decline, visual loss and eventually premature death (49). NCLs result due to mutations in at least one of the 14 different genes that have been identified and associated with NCL (50). The incidence rate of NCLs is estimated to be 1:56,000 – 1:67,000 (51) in the USA and 1:12,500 in the Scandinavian countries (52). The genes that are known to cause NCLs are – *CLN1/ PPT1, CLN2/TPP1, CLN3, CLN4/DNAJC5, CLN5, CLN6, CLN7/MFSD8, CLN8, CLN9, and CLN10* (53). Although there have been extensive studies done on the NCL proteins, the function of most of them is still unresolved (54). All NCL diseases have one factor in common; they all result in the accumulation of autofluorescent lipopigments within the lysosomes (50, 52, 55). Recently, a new nomenclature was proposed for the NCL diseases and the genes. This new nomenclature is according to the subtype of NCL followed by the age of onset of the subtype of NCL and electron microscopy (EM) findings (e.g., CLN3 disease, juvenile). To date there are 365 mutations known to cause NCL (53). The NCL genes are heterogeneous and consist of soluble lysosomal proteins (CLN1, CLN2, CLN5 and CLN10) as well as lysosomal transmembrane proteins (CLN3 and CLN7/MFSD) and some reside in the ER (CLN6 and CLN8) (54). Although NCL proteins are ubiquitously expressed in different cells and tissues, neuronal cells are most affected by defective NCL proteins and the manifestations of the NCL disease are first observed in these cells (53). It has become imperative to develop model organisms to study these diseases. Over the last few years several model organisms have been

developed with NCL diseases available. There are vertebrate and invertebrate models of NCL disease. Vertebrate models include sheep, dog, cow and mouse (http://www.ucl.ac.uk/ncl/animal.shtml). Invertebrate models include yeast (*S. cerevisiae* and *S. pombe*), *D. melanogaster* and *C. elegans* (56). Table 1.1 summarizes the different NCL genes, the resulting phenotype and the defects caused by mutations in these genes.

Table 1.1- Summary of different NCL genes

Gene	Disease/ Subtypes	Chromosomal location	Intracellular localization	Symptoms	Storage material	Function	Mutations known	Reference(s)
CLN1/PPT1	Classic infantile, late infantile, juvenile and adult	1p32	Lysosome (luminal)	Motor skill deterioration, seizures, visual and speech failure	SAP A and SAP D	Fatty acid hydrolase	61	(53, 57-64)
CLN2/TPP1	Classic infantile, infantile, and juvenile	11p15	Lysosome (luminal)	Mental and speech failure, ataxia, loss of vision, epilepsy and myoclonic jerks	Subunit c of mitochondrial ATP synthase	Tripeptidyl peptidase	89	(53) (65-70)
CLN3	Infantile, late infantile, juvenile, and adult	16p12.1	Lysosome (transmembrane)	Hallucinations, visual deterioration, loss of motor coordination	Subunit c of mitochondrial ATP synthase	Unknown	59	(53, 71-78)
CLN4/ DNAJC5	Adult	20q13.33	Secretory vesicles in neurons	Dementia, ataxia, seizures and myoclonus	Sap A and D or Subunit c of mitochondrial ATP synthase	Cysteine string protein alpha (CSP α)	-	(79-83)
CLN5	Infantile, late infantile variant, juvenile, and adult	13q21.1-q32	Lysosome (luminal)	Motor clumsiness, ataxia, loss of vision, seizures, mental retardation	Subunit c of mitochondrial ATP synthase	Unknown	33	(53) (84-92)

Gene	Disease/ Subtypes	Chromosomal location	Intracellular localization	Symptoms	Storage material	Function	Mutations known	Reference(s)
CLN6	Late infantile and Adult	15q21-q23	ER (Transmembrane)	Motor difficulties, seizures, ataxia, mental retardation, loss of vision	Subunit c of mitochondrial ATP synthase	Unknown	63	(65, 93)) (53) (93-96)
CLN7/ MFSD8	Late infantile variant and Juvenile	4q28.1-q28.2	Lysosome (transmembrane)	Developmental decadence, seizures, ataxia, vision failure, ambulatory problems	-	Lysosomal transporter	31	(53) (97-100)
CLN8	Variant late infantile and EPMR	8p23	ER (Transmembrane)	Ataxia, speech delay, myoclonus, loss of vision, seizures and mental decline	Subunit c of mitochondrial ATP synthase	Unknown	25	(53)(101-106)
CLN10/ CTSD	Congenital	11p15.5	Lysosome (luminal)	Post natal respiratory insufficiency and status epilepticus	-	Aspartyl protease	4	(53) (107-110)

CLN5

This thesis work focuses mainly on CLN5 protein and the role of N-glycosylation in the trafficking of CLN5 to the lysosome. Thus it is imperative to discuss CLN5 in more detail. The CLN5 disease has different forms – infantile, late infantile, juvenile and adult. It is also known as variable late infantile (vLINCL). The disease onset is usually between 4-7 years of age and the main symptom observed is motor clumsiness (84, 85) along with epileptic seizures, ataxia, mental and visual deterioration (84, 86, 87). Children usually go blind by the age of 7-10 and death occurs between the ages of 10-30 years (86, 95). Brain studies have shown cerebellar atrophy to be a notable characteristic of CLN5 disease (86). In CLN5 disease, late infantile and juvenile are associated with Fingerprint (FP) morphology of storage granules occasionally mixed with Curvilinear bodies (CL) or rectilinear bodies (RL) (94, 111). FP and granular osmiophilic deposits (GROD) lipopigments morphology are observed in the adult and infantile subtype of CLN5 (51, 112). Subunit c of mitochondrial ATP synthase is the main storage material that is seen to accumulate in lysosomes (88).

CLN5 maps to chromosome 13q21.1-q32 and consists of 407 amino acids (89-91). CLN5 is heavily glycosylated with a reported molecular weight of ~ 60 kDa (90, 113). One of the unusual characteristic of CLN5 protein is that it possesses a long signal peptide sequence (96 aa) (89). CLN5 does not share any homology with any other protein(s) and its function remains to be elucidated (90, 91). There has been speculation of the intracellular localization of CLN5 but it is generally accepted to be a soluble lysosomal glycoprotein (89, 90, 114). Lysosomal targeting of CLN5 follows M6PR pathway, although M6PR independent pathways have also been shown to be utilized by CLN5 to reach the lysosomes (81, 89, 115).

To date, there are 33 mutations known for *CLN5* gene (53). All the missense mutations studied so far (p.R112H, p.R112P, p.D279N, and p.W379C) cause retention of CLN5 in ER possibly due to misfolding or defect in trafficking (89, 116). Although *CLN5* mutations were initially reported mainly in Finnish populations, it is now clear that *CLN5* mutations have a worldwide distribution with mutations reported in over 15 different countries (51). Even though the function of CLN5 is not known, it has been reported to interact with PPT1/CLN1, CLN6 and CLN8 in ER (117) and with TPP1/CLN2 and CLN3 in the lysosomes (118). It has also been implicated to play a role in lysosomal receptor recycling via the recruitment of retromer (113).

Mutations in the N-glycosylation sites have been known to have a drastic effect on other NCL proteins (CLN1 and CLN2) by affecting their activity and/or localization. CLN2/TPP1 has five N-glycosylation sites of which one of them (Asn286) has been shown to be important for targeting to the lysosome (119). The role of N-glycosylation in proper trafficking of CLN5 to the lysosome has not been studied. Such a study is interesting as CLN5 has eight potential N-glycosylation sites and to elucidate which ones are important will be essential in characterizing it.

References

1. **Rothman JE, Wieland FT** 1996 Protein sorting by transport vesicles. Science 272:227-234

2. **Pfeffer SR, Rothman JE** 1987 Biosynthetic protein transport and sorting by the endoplasmic reticulum and Golgi. Annu Rev Biochem 56:829-852

3. **Watanabe R, Riezman H** 2004 Differential ER exit in yeast and mammalian cells. Curr Opin Cell Biol 16:350-355

4. **Lee MC, Miller EA, Goldberg J, Orci L, Schekman R** 2004 Bi-directional protein transport between the ER and Golgi. Annu Rev Cell Dev Biol 20:87-123

5. **Thomas G** 2002 Furin at the cutting edge: from protein traffic to embryogenesis and disease. Nat Rev Mol Cell Biol 3:753-766

6. **Martinez-Menarguez JA, Geuze HJ, Slot JW, Klumperman J** 1999 Vesicular tubular clusters between the ER and Golgi mediate concentration of soluble secretory proteins by exclusion from COPI-coated vesicles. Cell 98:81-90

7. **Barlowe C** 2003 Signals for COPII-dependent export from the ER: what's the ticket out? Trends Cell Biol 13:295-300

8. **Bonifacino JS, Glick BS** 2004 The mechanisms of vesicle budding and fusion. Cell 116:153-166

9. **Giraudo CG, Maccioni HJ** 2003 Endoplasmic reticulum export of glycosyltransferases depends on interaction of a cytoplasmic dibasic motif with Sar1. Mol Biol Cell 14:3753-3766

10. **Pelham HR** 1989 Control of protein exit from the endoplasmic reticulum. Annu Rev Cell Biol 5:1-23

11. **Nilsson T, Warren G** 1994 Retention and retrieval in the endoplasmic reticulum and the Golgi apparatus. Curr Opin Cell Biol 6:517-521

12. **Appenzeller-Herzog C, Hauri HP** 2006 The ER-Golgi intermediate compartment (ERGIC): in search of its identity and function. J Cell Sci 119:2173-2183

13. **Farquhar MG, Palade GE** 1981 The Golgi apparatus (complex)-(1954-1981)-from artifact to center stage. J Cell Biol 91:77s-103s

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

15. **Ghosh P, Dahms NM, Kornfeld S** 2003 Mannose 6-phosphate receptors: new twists in the tale. Nat Rev Mol Cell Biol 4:202-212

16. **Huotari J, Helenius A** 2011 Endosome maturation. EMBO J 30:3481-3500

17. **Rink J, Ghigo E, Kalaidzidis Y, Zerial M** 2005 Rab conversion as a mechanism of progression from early to late endosomes. Cell 122:735-749

18. **Poteryaev D, Datta S, Ackema K, Zerial M, Spang A** 2010 Identification of the switch in early-to-late endosome transition. Cell 141:497-508

19. **Kirchhausen T, Bonifacino JS, Riezman H** 1997 Linking cargo to vesicle formation: receptor tail interactions with coat proteins. Curr Opin Cell Biol 9:488-495

20. **Ohno H, Aguilar RC, Yeh D, Taura D, Saito T, Bonifacino JS** 1998 The medium subunits of adaptor complexes recognize distinct but overlapping sets of tyrosine-based sorting signals. J Biol Chem 273:25915-25921

21. **Ohno H, Fournier MC, Poy G, Bonifacino JS** 1996 Structural determinants of interaction of tyrosine-based sorting signals with the adaptor medium chains. J Biol Chem 271:29009-29015

22. **Ohno H, Stewart J, Fournier MC, Bosshart H, Rhee I, Miyatake S, Saito T, Gallusser A, Kirchhausen T, Bonifacino JS** 1995 Interaction of tyrosine-based sorting signals with clathrin-associated proteins. Science 269:1872-1875

23. **Boll W, Ohno H, Songyang Z, Rapoport I, Cantley LC, Bonifacino JS, Kirchhausen T** 1996 Sequence requirements for the recognition of tyrosine-based endocytic signals by clathrin AP-2 complexes. EMBO J 15:5789-5795

24. **Rapoport I, Chen YC, Cupers P, Shoelson SE, Kirchhausen T** 1998 Dileucine-based sorting signals bind to the beta chain of AP-1 at a site distinct and regulated differently from the tyrosine-based motif-binding site. EMBO J 17:2148-2155

25. **Le Borgne R, Hoflack B** 1998 Mechanisms of protein sorting and coat assembly: insights from the clathrin-coated vesicle pathway. Curr Opin Cell Biol 10:499-503

26. **Helenius A, Aebi M** 2001 Intracellular functions of N-linked glycans. Science 291:2364-2369

27. **Lazzarino DA, Gabel CA** 1989 Mannose processing is an important determinant in the assembly of phosphorylated high mannose-type oligosaccharides. J Biol Chem 264:5015-5023

28. **Kornfeld R, Bao M, Brewer K, Noll C, Canfield W** 1999 Molecular cloning and functional expression of two splice forms of human N-acetylglucosamine-1-phosphodiester alpha-N-acetylglucosaminidase. J Biol Chem 274:32778-32785

29. **Rohrer J, Kornfeld R** 2001 Lysosomal hydrolase mannose 6-phosphate uncovering enzyme resides in the trans-Golgi network. Mol Biol Cell 12:1623-1631

30. **Braulke T, Gartung C, Hasilik A, von Figura K** 1987 Is movement of mannose 6-phosphate-specific receptor triggered by binding of lysosomal enzymes? J Cell Biol 104:1735-1742

31. **Breuer P, Korner C, Boker C, Herzog A, Pohlmann R, Braulke T** 1997 Serine phosphorylation site of the 46-kDa mannose 6-phosphate receptor is required for transport to the plasma membrane in Madin-Darby canine kidney and mouse fibroblast cells. Mol Biol Cell 8:567-576

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

33. **Lefrancois S, Zeng J, Hassan AJ, Canuel M, Morales CR** 2003 The lysosomal trafficking of sphingolipid activator proteins (SAPs) is mediated by sortilin. EMBO J 22:6430-6437

34. **Ni X, Morales CR** 2006 The lysosomal trafficking of acid sphingomyelinase is mediated by sortilin and mannose 6-phosphate receptor. Traffic 7:889-902

35. **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 131:770-783

36. **Traub LM, Ostrom JA, Kornfeld S** 1993 Biochemical dissection of AP-1 recruitment onto Golgi membranes. J Cell Biol 123:561-573

37. **Seaman MN, Sowerby PJ, Robinson MS** 1996 Cytosolic and membrane-associated proteins involved in the recruitment of AP-1 adaptors onto the trans-Golgi network. J Biol Chem 271:25446-25451

38. **Meyer C, Zizioli D, Lausmann S, Eskelinen EL, Hamann J, Saftig P, von Figura K, Schu P** 2000 mu1A-adaptin-deficient mice: lethality, loss of AP-1 binding and rerouting of mannose 6-phosphate receptors. EMBO J 19:2193-2203

39. **Journet A, Chapel A, Kieffer S, Roux F, Garin J** 2002 Proteomic analysis of human lysosomes: application to monocytic and breast cancer cells. Proteomics 2:1026-1040

40. **Wraith JE** 2002 Lysosomal disorders. Semin Neonatol 7:75-83

41. **Blanz J, Groth J, Zachos C, Wehling C, Saftig P, Schwake M** 2010 Disease-causing mutations within the lysosomal integral membrane protein type 2 (LIMP-2) reveal the nature of binding to its ligand beta-glucocerebrosidase. Hum Mol Genet 19:563-572

42. **Brett CL, Kallay L, Hua Z, Green R, Chyou A, Zhang Y, Graham TR, Donowitz M, Rao R** 2011 Genome-wide analysis reveals the vacuolar pH-stat of Saccharomyces cerevisiae. PLoS One 6:e17619

43. **Marshansky V, Futai M** 2008 The V-type H+-ATPase in vesicular trafficking: targeting, regulation and function. Curr Opin Cell Biol 20:415-426

44. **Weimer RM, Jorgensen EM** 2003 Controversies in synaptic vesicle exocytosis. J Cell Sci 116:3661-3666

45. **Sobota JA, Back N, Eipper BA, Mains RE** 2009 Inhibitors of the V0 subunit of the vacuolar H+-ATPase prevent segregation of lysosomal- and secretory-pathway proteins. J Cell Sci 122:3542-3553

46. **Desnick RJ, Schuchman EH** 2002 Enzyme replacement and enhancement therapies: lessons from lysosomal disorders. Nat Rev Genet 3:954-966

47. **Padilla-Lopez S, Pearce DA** 2006 Saccharomyces cerevisiae lacking Btn1p modulate vacuolar ATPase activity to regulate pH imbalance in the vacuole. J Biol Chem 281:10273-10280

48. **Sardiello M, Palmieri M, di Ronza A, Medina DL, Valenza M, Gennarino VA, Di Malta C, Donaudy F, Embrione V, Polishchuk RS, Banfi S, Parenti G, Cattaneo E, Ballabio A** 2009 A gene network regulating lysosomal biogenesis and function. Science 325:473-477

49. **Haltia M** 2003 The neuronal ceroid-lipofuscinoses. J Neuropathol Exp Neurol 62:1-13

50. **Kohlschutter A, Schulz A** 2009 Towards understanding the neuronal ceroid lipofuscinoses. Brain Dev 31:499-502

51. **Xin W, Mullen TE, Kiely R, Min J, Feng X, Cao Y, O'Malley L, Shen Y, Chu-Shore C, Mole SE, Goebel HH, Sims K** 2010 CLN5 mutations are frequent in juvenile and late-onset non-Finnish patients with NCL. Neurology 74:565-571

52. **Santavuori P** 1988 Neuronal ceroid-lipofuscinoses in childhood. Brain Dev 10:80-83

53. **Kousi M, Lehesjoki AE, Mole SE** 2012 Update of the mutation spectrum and clinical correlations of over 360 mutations in eight genes that underlie the neuronal ceroid lipofuscinoses. Hum Mutat 33:42-63

54. **Jalanko A, Braulke T** 2009 Neuronal ceroid lipofuscinoses. Biochim Biophys Acta 1793:697-709

55. **Haltia M** 2006 The neuronal ceroid-lipofuscinoses: from past to present. Biochim Biophys Acta 1762:850-856

56. **Phillips SN, Muzaffar N, Codlin S, Korey CA, Taschner PE, de Voer G, Mole SE, Pearce DA** 2006 Characterizing pathogenic processes in Batten disease: use of small eukaryotic model systems. Biochim Biophys Acta 1762:906-919

57. **Das AK, Becerra CH, Yi W, Lu JY, Siakotos AN, Wisniewski KE, Hofmann SL** 1998 Molecular genetics of palmitoyl-protein thioesterase deficiency in the U.S. J Clin Invest 102:361-370

58. **Ramadan H, Al-Din AS, Ismail A, Balen F, Varma A, Twomey A, Watts R, Jackson M, Anderson G, Green E, Mole SE** 2007 Adult neuronal ceroid lipofuscinosis caused by deficiency in palmitoyl protein thioesterase 1. Neurology 68:387-388

59. **Santavuori P, Lauronen L, Kirveskari K, Aberg L, Sainio K** 2000 Neuronal ceroid lipofuscinoses in childhood. Suppl Clin Neurophysiol 53:443-451

60. **Wisniewski KE, Zhong N, Kaczmarski W, Kaczmarski A, Kida E, Brown WT, Schwarz KO, Lazzarini AM, Rubin AJ, Stenroos ES, Johnson WG, Wisniewski TM** 1998 Compound heterozygous genotype is associated with protracted juvenile neuronal ceroid lipofuscinosis. Ann Neurol 43:106-110

61. **Vesa J, Hellsten E, Verkruyse LA, Camp LA, Rapola J, Santavuori P, Hofmann SL, Peltonen L** 1995 Mutations in the palmitoyl protein thioesterase gene causing infantile neuronal ceroid lipofuscinosis. Nature 376:584-587

62. **Jarvela I, Schleutker J, Haataja L, Santavuori P, Puhakka L, Manninen T, Palotie A, Sandkuijl LA, Renlund M, White R** 1991 Infantile form of neuronal ceroid lipofuscinosis (CLN1) maps to the short arm of chromosome 1. Genomics 9:170-173

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

64. **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

65. **Sharp JD, Wheeler RB, Lake BD, Savukoski M, Jarvela IE, Peltonen L, Gardiner RM, Williams RE** 1997 Loci for classical and a variant late infantile neuronal ceroid lipofuscinosis map to chromosomes 11p15 and 15q21-23. Hum Mol Genet 6:591-595

66. **Sleat DE, Donnelly RJ, Lackland H, Liu CG, Sohar I, Pullarkat RK, Lobel P** 1997 Association of mutations in a lysosomal protein with classical late-infantile neuronal ceroid lipofuscinosis. Science 277:1802-1805

67. **Steinfeld R, Steinke HB, Isbrandt D, Kohlschutter A, Gartner J** 2004 Mutations in classical late infantile neuronal ceroid lipofuscinosis disrupt transport of tripeptidyl-peptidase I to lysosomes. Hum Mol Genet 13:2483-2491

68. **Vines D, Warburton MJ** 1998 Purification and characterisation of a tripeptidyl aminopeptidase I from rat spleen. Biochim Biophys Acta 1384:233-242

69. **Warburton MJ, Bernardini F** 2000 Tripeptidyl-peptidase I deficiency in classical late-infantile neuronal ceroid lipofuscinosis brain tissue. Evidence for defective peptidase rather than proteinase activity. J Inherit Metab Dis 23:145-154

70. **Ezaki J, Takeda-Ezaki M, Oda K, Kominami E** 2000 Characterization of endopeptidase activity of tripeptidyl peptidase-I/CLN2 protein which is deficient in classical late infantile neuronal ceroid lipofuscinosis. Biochem Biophys Res Commun 268:904-908

71. **Jarvela I, Autti T, Lamminranta S, Aberg L, Raininko R, Santavuori P** 1997 Clinical and magnetic resonance imaging findings in Batten disease: analysis of the major mutation (1.02-kb deletion). Ann Neurol 42:799-802

72. **Bensaoula T, Shibuya H, Katz ML, Smith JE, Johnson GS, John SK, Milam AH** 2000 Histopathologic and immunocytochemical analysis of the retina and ocular tissues in Batten disease. Ophthalmology 107:1746-1753

73. **Eksandh LB, Ponjavic VB, Munroe PB, Eiberg HE, Uvebrant PE, Ehinger BE, Mole SE, Andreasson S** 2000 Full-field ERG in patients with Batten/Spielmeyer-Vogt disease caused by mutations in the CLN3 gene. Ophthalmic Genet 21:69-77

74. **Palmer DN, Fearnley IM, Walker JE, Hall NA, Lake BD, Wolfe LS, Haltia M, Martinus RD, Jolly RD** 1992 Mitochondrial ATP synthase subunit c storage in the ceroid-lipofuscinoses (Batten disease). Am J Med Genet 42:561-567

75. **Ezaki J, Takeda-Ezaki M, Koike M, Ohsawa Y, Taka H, Mineki R, Murayama K, Uchiyama Y, Ueno T, Kominami E** 2003 Characterization of Cln3p, the gene product responsible for juvenile neuronal ceroid lipofuscinosis, as a lysosomal integral membrane glycoprotein. J Neurochem 87:1296-1308

76. **Kyttala A, Ihrke G, Vesa J, Schell MJ, Luzio JP** 2004 Two motifs target Batten disease protein CLN3 to lysosomes in transfected nonneuronal and neuronal cells. Mol Biol Cell 15:1313-1323

77. **Phillips SN, Benedict JW, Weimer JM, Pearce DA** 2005 CLN3, the protein associated with batten disease: structure, function and localization. J Neurosci Res 79:573-583

78. **Haskell RE, Derksen TA, Davidson BL** 1999 Intracellular trafficking of the JNCL protein CLN3. Mol Genet Metab 66:253-260

79. **Berkovic SF, Andermann F, Andermann E, Carpenter S, Wolfe L** 1988 Kufs disease: clinical features and forms. Am J Med Genet Suppl 5:105-109

80. **Nijssen PC, Brusse E, Leyten AC, Martin JJ, Teepen JL, Roos RA** 2002 Autosomal dominant adult neuronal ceroid lipofuscinosis: parkinsonism due to both striatal and nigral dysfunction. Mov Disord 17:482-487

81. **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 8:1708-1718

82. **Nijssen PC, Ceuterick C, van Diggelen OP, Elleder M, Martin JJ, Teepen JL, Tyynela J, Roos RA** 2003 Autosomal dominant adult neuronal ceroid lipofuscinosis: a novel form of NCL with granular osmiophilic deposits without palmitoyl protein thioesterase 1 deficiency. Brain Pathol 13:574-581

83. **Hall NA, Lake BD, Dewji NN, Patrick AD** 1991 Lysosomal storage of subunit c of mitochondrial ATP synthase in Batten's disease (ceroid-lipofuscinosis). Biochem J 275 (Pt 1):269-272

84. **Santavuori P, Rapola J, Sainio K, Raitta C** 1982 A variant of Jansky-Bielschowsky disease. Neuropediatrics 13:135-141

85. **Santavuori P, Rapola J, Raininko R, Autti T, Lappi M, Nuutila A, Launes J, Sainio K** 1993 Early juvenile neuronal ceroid-lipofuscinosis or variant Jansky-Bielschowsky disease: diagnostic criteria and nomenclature. J Inherit Metab Dis 16:230-232

86. **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 55:579-581

87. **Santavuori P, Rapola J, Nuutila A, Raininko R, Lappi M, Launes J, Herva R, Sainio K** 1991 The spectrum of Jansky-Bielschowsky disease. Neuropediatrics 22:92-96

88. **Tyynela J, Suopanki J, Santavuori P, Baumann M, Haltia M** 1997 Variant late infantile neuronal ceroid-lipofuscinosis: pathology and biochemistry. J Neuropathol Exp Neurol 56:369-375

89. **Schmiedt ML, Bessa C, Heine C, Ribeiro MG, Jalanko A, Kyttala A** 2010 The neuronal ceroid lipofuscinosis protein CLN5: new insights into cellular maturation, transport, and consequences of mutations. Hum Mutat 31:356-365

90. **Isosomppi J, Vesa J, Jalanko A, Peltonen L** 2002 Lysosomal localization of the neuronal ceroid lipofuscinosis CLN5 protein. Hum Mol Genet 11:885-891

91. **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 19:286-288

92. **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 16:29-40

93. **Gao H, Boustany RM, Espinola JA, Cotman SL, Srinidhi L, Antonellis KA, Gillis T, Qin X, Liu S, Donahue LR, Bronson RT, Faust JR, Stout D, Haines JL, Lerner TJ, MacDonald ME** 2002 Mutations in a novel CLN6-encoded transmembrane protein cause variant neuronal ceroid lipofuscinosis in man and mouse. Am J Hum Genet 70:324-335

94. **Mole SE, Williams RE, Goebel HH** 2005 Correlations between genotype, ultrastructural morphology and clinical phenotype in the neuronal ceroid lipofuscinoses. Neurogenetics 6:107-126

95. **Moore SJ, Buckley DJ, MacMillan A, Marshall HD, Steele L, Ray PN, Nawaz Z, Baskin B, Frecker M, Carr SM, Ives E, Parfrey PS** 2008 The clinical and genetic epidemiology of neuronal ceroid lipofuscinosis in Newfoundland. Clin Genet 74:213-222

96. **Sharp JD, Wheeler RB, Parker KA, Gardiner RM, Williams RE, Mole SE** 2003 Spectrum of CLN6 mutations in variant late infantile neuronal ceroid lipofuscinosis. Hum Mutat 22:35-42

97. **Kousi M, Siintola E, Dvorakova L, Vlaskova H, Turnbull J, Topcu M, Yuksel D, Gokben S, Minassian BA, Elleder M, Mole SE, Lehesjoki AE** 2009 Mutations in CLN7/MFSD8 are a common cause of variant late-infantile neuronal ceroid lipofuscinosis. Brain 132:810-819

98. **Topcu M, Tan H, Yalnizoglu D, Usubutun A, Saatci I, Aynaci M, Anlar B, Topaloglu H, Turanli G, Kose G, Aysun S** 2004 Evaluation of 36 patients from Turkey with neuronal ceroid lipofuscinosis: clinical, neurophysiological, neuroradiological and histopathologic studies. Turk J Pediatr 46:1-10

99. **Siintola E, Topcu M, Aula N, Lohi H, Minassian BA, Paterson AD, Liu XQ, Wilson C, Lahtinen U, Anttonen AK, Lehesjoki AE** 2007 The novel neuronal ceroid lipofuscinosis gene MFSD8 encodes a putative lysosomal transporter. Am J Hum Genet 81:136-146

100. **Sharifi A, Kousi M, Sagne C, Bellenchi GC, Morel L, Darmon M, Hulkova H, Ruivo R, Debacker C, El Mestikawy S, Elleder M, Lehesjoki AE, Jalanko A, Gasnier B, Kyttala A** 2010 Expression and lysosomal targeting of CLN7, a major facilitator superfamily transporter associated with variant late-infantile neuronal ceroid lipofuscinosis. Hum Mol Genet 19:4497-4514

101. **Herva R, Tyynela J, Hirvasniemi A, Syrjakallio-Ylitalo M, Haltia M** 2000 Northern epilepsy: a novel form of neuronal ceroid-lipofuscinosis. Brain Pathol 10:215-222

102. **Hirvasniemi A, Lang H, Lehesjoki AE, Leisti J** 1994 Northern epilepsy syndrome: an inherited childhood onset epilepsy with associated mental deterioration. J Med Genet 31:177-182

103. **Mitchell WA, Wheeler RB, Sharp JD, Bate SL, Gardiner RM, Ranta US, Lonka L, Williams RE, Lehesjoki AE, Mole SE** 2001 Turkish variant late infantile neuronal ceroid lipofuscinosis (CLN7) may be allelic to CLN8. Eur J Paediatr Neurol 5 Suppl A:21-27

104. **Ranta S, Topcu M, Tegelberg S, Tan H, Ustubutun A, Saatci I, Dufke A, Enders H, Pohl K, Alembik Y, Mitchell WA, Mole SE, Lehesjoki AE** 2004 Variant late infantile neuronal ceroid lipofuscinosis in a subset of Turkish patients is allelic to Northern epilepsy. Hum Mutat 23:300-305

105. **Ranta S, Zhang Y, Ross B, Lonka L, Takkunen E, Messer A, Sharp J, Wheeler R, Kusumi K, Mole S, Liu W, Soares MB, Bonaldo MF, Hirvasniemi A, de la Chapelle A, Gilliam TC, Lehesjoki AE** 1999 The neuronal ceroid lipofuscinoses in human EPMR and mnd mutant mice are associated with mutations in CLN8. Nat Genet 23:233-236

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

107. **Fritchie K, Siintola E, Armao D, Lehesjoki AE, Marino T, Powell C, Tennison M, Booker JM, Koch S, Partanen S, Suzuki K, Tyynela J, Thorne LB** 2009 Novel mutation and the first prenatal screening of cathepsin D deficiency (CLN10). Acta Neuropathol 117:201-208

108. **Sandbank U** 1968 Congenital amaurotic idiocy. Pathol Eur 3:226-229

109. **Scarborough PE, Dunn BM** 1994 Redesign of the substrate specificity of human cathepsin D: the dominant role of position 287 in the S2 subsite. Protein Eng 7:495-502

110. **Zaidi N, Maurer A, Nieke S, Kalbacher H** 2008 Cathepsin D: a cellular roadmap. Biochem Biophys Res Commun 376:5-9

111. **Pineda-Trujillo N, Cornejo W, Carrizosa J, Wheeler RB, Munera S, Valencia A, Agudelo-Arango J, Cogollo A, Anderson G, Bedoya G, Mole SE, Ruiz-Linares A** 2005 A CLN5 mutation causing an atypical neuronal ceroid lipofuscinosis of juvenile onset. Neurology 64:740-742

112. **Cismondi IA, Cannelli N, Aiello C, Santorelli FM, Kohan R, Oller Ramirez AM, Halac IN** 2008 Gene symbol: CLN5. Disease: Neuronal Ceroid Lipofuscinosis, Finnish Variant. Hum Genet 123:554

113. **Mamo A, Jules F, Dumaresq-Doiron K, Costantino S, Lefrancois S** 2012 The Role of Ceroid Lipofuscinosis Neuronal Protein 5 (CLN5) in Endosomal Sorting. Mol Cell Biol 32:1855-1866

114. **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 16:29-40

115. **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 5:3966-3978

116. **Lebrun AH, Storch S, Ruschendorf F, Schmiedt ML, Kyttala A, Mole SE, Kitzmuller C, Saar K, Mewasingh LD, Boda V, Kohlschutter A, Ullrich K, Braulke T, Schulz A** 2009 Retention of lysosomal protein CLN5 in the endoplasmic reticulum causes neuronal ceroid lipofuscinosis in Asian sibship. Hum Mutat 30:E651-61

117. **Lyly A, von Schantz C, Heine C, Schmiedt ML, Sipila T, Jalanko A, Kyttala A** 2009 Novel interactions of CLN5 support molecular networking between Neuronal Ceroid Lipofuscinosis proteins. BMC Cell Biol 10:83

118. **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 13:2410-2420

119. **Wujek P, Kida E, Walus M, Wisniewski KE, Golabek AA** 2004 N-glycosylation is crucial for folding, trafficking, and stability of human tripeptidyl-peptidase I. J Biol Chem 279:12827-12839

CHAPTER 2

ROLE OF N-GLYCOSYLATION IN TRAFFICKING AND STABILITY OF HUMAN CLN5

ABSTRACT

CLN5 is a soluble lysosomal protein with unknown function. Mutations in *CLN5* lead to neuronal-ceroid lipofuscinosis (NCL), a group of inherited neurodegenerative lysosomal storage disorders that mainly affect children. NCLs have distinct clinical symptoms that include mental retardation, loss of vision, epilepsy, and impediment of motor capabilities. Buildup of autofluorescent lipopigments called lipofuscin in the lysosome is observed with subunit c of mitochondrial ATP synthase being the main storage material. CLN5 is a heavily glycosylated protein with eight potential N-glycosylation sites based on the Asn-X-Ser/Thr consensus sequence. It has been shown to utilize the mannose 6-phosphate (M6P) pathway to reach the lysosomes. By mutating asparagine to glutamine in the N-glycosylation consensus sequences, we showed that all eight putative N-glycosylation sites are used *in vivo*. From localization studies, we found that certain N-glycosylation sites are important for protein folding and trafficking as mutating these sites (N179Q, N252Q, N304Q, and N320Q) lead to the retention of CLN5 in the ER. One particular site, N401, is essential for trafficking out from the Golgi, as CLN5 with N401Q mutation mislocalized to the Golgi. We also analyzed several known patient mutations that result in the mislocalization of CLN5 similar to the N-glycosylation mutants.

INTRODUCTION

The Neuronal ceroid lipofuscinosis (NCLs), also known as Batten disease, are a group of progressive neurodegenerative disorders mainly affecting children. The NCLs are characterized by mental retardation, impediment of motor capabilities, loss of vision, epileptic seizures, and eventually causing death. Depending on the onset, progression and the neurological symptoms, there are several forms of NCL that have been identified as infantile (INCL), late infantile (LINCL), juvenile (JNCL) and adult (ANCL) (1). The incidence of this disease is 1:100,000 worldwide (2) but it is more common in the USA where the incidence rate is between 1:56,000 to 1:67,000 (3) and even higher in Scandinavian countries with the incidence rate being 1:12,500 (2). There are at least fourteen different established disorders falling under NCL (4) but the exact function of most of these genes (*CLN1-CLN14*) is not well established (1). NCL is classified as a lysosomal storage disorder based on the buildup of autofluorescent lipopigments called lipofuscin in lysosomes of neurons as well as other cell types (2, 5, 6). The main storage material usually seen accumulated is the subunit c of mitochondrial ATP synthase (7) and Saposin A and D (8). In mouse models of NCL disease some secondary lipids such as glycosphingolipids, dolichol, bis (monoacylglycero) phosphate or cholesterol have also been reported to accumulate (9). Although NCL proteins are ubiquitously expressed, neuronal cells are most affected by dysfunctional NCL proteins (10).

CLN5 mutations were initially reported to be limited to Finnish and other Northern European populations, but recently it has been reported to be present in Columbian, Portuguese, Pakistani, Afghan and Italian population, making the disease’s ethnic distribution

more diverse (3). The exact function of this protein is not known and it does not share homology with any other protein(s) (11, 12). CLN5 protein consists of 407 amino acids with a signal sequence of ~96 amino acids. Although *in silico* analysis suggests it has a transmembrane domain (11, 12), it has been shown to be a soluble protein (13) and is reported to localize to the endosomal-lysosomal compartment (13, 14). Transient over-expression of CLN5 demonstrated that the protein has a molecular weight of ~60 kDa and is heavily glycosylated (14, 15). Stable expression of CLN5 in a SH-SY5Y neuronal cell line showed that CLN5 can be detected as a 60 kDa proform and 50 kDa mature form before reaching the lysosomes (16). Although not much is known about CLN5, it has been reported to interact with other NCL proteins – CLN2 and CLN3 (12) as well as CLN1/PPT1, CLN6, and CLN8 (17). Interaction of CLN5 with CLN2 is reported to take place in the late endocytic compartment and it has been suggested that CLN5 may be a connecting link between the different NCL proteins (17). A recent study reported CLN5 to be integral for the recruitment of retromer, which in turn was responsible for the sorting and recycling of lysosomal receptors (15) .

Post-translational modifications play an important role in the maturation, folding and trafficking of many proteins. One of the main modifications proteins undergo during and after translation is glycosylation. N-glycosylation is the addition of oligosaccharides onto an asparagine (Asn) residue. The addition of oligosaccharide chain to Asn is not random and usually follows a consensus sequence of N-X-S/T where X is any amino acid apart from proline. The addition of core sugar structure takes place in the ER and further modifications to the sugar (trimming or additions) take place in the Golgi (e.g., phosphorylation of mannose) (18-20). One of the best characterized role of N-glycosylation is the targeting of acid hydrolases to the

lysosomes. The Mannose-6-phosphate (Man-6-P) residue on the acid hydrolases is recognized by mannose 6-phosphate receptors (M6PR) in the trans-Golgi network which carry their cargo to the late endosomes/lysosomes.

CLN5 is a highly glycosylated protein with almost 40% of its molecular weight attributed to glycosylation. This has been demonstrated by the digestion of CLN5 with Endo H (cleaves mostly high mannose sugars), PNGase F (cleaves N-linked glycosylation) and also by tunicamycin treatment (prevents N-glycosylation) (14, 16) which results in the reduction of size of the protein from ~60 kDa to ~35 kDa. Based on the consensus sequence N-X-S/T, CLN5 has eight potential N-glycosylation sites at Asn179, 192, 227, 252, 304, 320, 330, and 401 (12, 14). Proteins that are destined to the lysosome normally contain one or more M6P residues on their N-linked oligosaccharides, which are recognized by M6PR and escorted to the lysosomes. Mass spectrometry analysis has predicted that human CLN5 potentially contains three such M6P sites at Asn 320, 330, and 401 (21). M6PR bind to M6P residues and transport the cargo protein to the endosome, where the M6PR detach from their cargo due to the acidic environment and are recycled to the (TGN) and are ready to bind cargo again (20). The CLN5 protein is then able to reach the lysosome from the endosome. Even though CLN5 has been predicted to contain three M6P residues, it has been reported that CLN5 is capable of utilizing M6PR independent pathway(s) in addition to the classical M6PR pathway to reach the lysosomes (16). Tripeptidyl-peptidase I (TPPI) is a good example of a lysosomal enzyme which is targeted to the lysosomes via the MPR dependent pathway (18, 22). It has five N-glycosylation sites, of which one of them (Asn286) has been shown to be the most important for the transport and activity of TPPI (18).

Here we studied the role of N-glycosylation on the trafficking of CLN5 protein. We utilize site directed mutagenesis to create mutants for each of the eight predicted N-glycosylation sites and look at the change in size and localization of each of these mutants using western blotting and immunofluorescence. We show that all the eight potential N-glycosylation sites are used *in vivo* and that the loss of N-glycosylation at certain sites leads to CLN5 being mislocalized either to the ER or the Golgi, whereas other mutants do not mislocalize and are able to reach the lysosome. This study sheds light on the importance of N-glycosylation on the transport of human CLN5 protein. The progress in understanding CLN5 has been hindered since not much is known about this protein. Here we show importance of some of the basic modifications (N-glycosylation) on CLN5 and how changing these modifications have a drastic effect on the localization of CLN5.

MATERIALS AND METHODS

Reagents

Cell culture media and reagents were purchased from Hyclone. Chemiluminescent HRP substrate kit was purchased from Millipore. Endoglycosidase H (Endo H_f), peptide-N-glycosidase F (PNGase F), O-glycosidase, T4 Polynucleotide kinase (PNK), Quick ligase, colorplus protein marker and restriction enzymes were from New England Biolabs (NEB). The TransIT-LT1 transfection reagent was from Mirus life technologies. The phusion PCR kit and chemically competent DH5-α *E.coli* cells were from NEB. Cycloheximide and 37% paraformaldehyde was

purchased from Fisher scientific. Tunicamycin was from Enzo life sciences. Rab5 mutant plasmid eGFP-Rab5A Q79L was purchased from Addgene (Addgene plasmid 28046) (23).

Antibodies

Mouse monoclonal antibody (α-Myc) producing hybridoma cell line 9E10 was purchased from ATCC (CRL 1729). Rabbit polyclonal antibodies used in this study were Calnexin for ER (Genscript), Grasp65 for Golgi (Affinity bioreagents), Lamp1, Lamp2 for Lysosomes, EEA1 for early Endosomes, and goat anti VPS35/MEM-3 for Retromer (all from Pierce). HRP-conjugated secondary antibodies for western blotting were purchased from Jackson laboratories. Secondary antibodies conjugated to Alexa Flour 488, 546, 647, and 633 were from Molecular probes. Rabbit α-CLN5 Ab (IRmII-4) was kindly provided by Dr. Kyttala (13). Mouse α-Lamp1 (H4A3) was purchased from the Developmental Studies Hybridoma Bank (Iowa City, IA).

Site-directed mutagenesis

The cDNA encoding wild type CLN5 (Purchased from Gene Copoeia, MGC: 184226) was cloned into pcDNA3.1/myc-His(-)A between EcoRI and BamHI restriction sites. This was used as the template for preparing individual N-glycosylation mutants by changing the codon for Asn to Gln within the consensus sequence for N-glycosylation Asn-X-Ser/Thr. Mutated cDNAs were generated using phusion site-directed mutagenesis kit. cDNAs containing single mutation for the N-glycosylation sites served as templates for creating multiple N-glycosylation mutants. In

these multiple mutants also, the codon for Asn was replaced with Gln. The mutations were confirmed by sequence analysis. The primer sequences used to generate different N-glycosylation mutants are shown in table 2.1.

Cell culture and transfections

Cells were grown and maintained in Dulbecco's modified eagle medium (HeLa cells ATCC CCL-2) supplemented with 10% fetal bovine serum (FBS), glutamax, HEPES and gentamycin at 37° C in a humidified incubator with 5% CO_2. 24 h before transfection, the cells were seeded on a 6-well culture plate at ~30% confluency. Cells were transiently transfected using Mirus Trans IT-LT1 transfection reagent according to manufacturer's recommendation. 24 h after transfection cells were collected and the cell pellet was either used immediately for western blotting or stored at -80° C until use.

SDS-PAGE and western blotting

Cells were lysed using RIPA lysis buffer (50 mM Tris pH 8, 150 mM NaCl, 1% NP40, 0.1% SDS, 0.5% sodium deoxycholate and protease inhibitor mix) on ice for 30 min. The clear lysate was collected after spinning at 13000 rpm for 5 min and sample buffer was added. For media samples, a 60% ammonium sulfate precipitation was carried out to precipitate proteins from serum free media. The samples were boiled for 5 min and loaded onto 10% SDS-PAGE. The electrophoretically separated proteins were then transferred onto a PVDF membrane. The

membrane was blocked with 5% milk in TBST (20 mM Tris-HCl pH 7.4, 150 mM NaCl and 0.05% Tween20) and incubated with primary antibody for 1.5 h at room temperature or overnight at 4° C. The membrane was then washed with TBST and incubated with the peroxidase-conjugated secondary antibody diluted to 1:50,000 in TBST for 45 min at room temperature and developed using Chemiluminescent HRP substrate. The membrane was imaged using G box imager from Syngene.

Immunofluorescence microscopy

Cells were grown on circular glass coverslips and transiently transfected as described before. The cells were treated with cycloheximide (50 µg/ ml) for 2 h before fixing with 4% formaldehyde for 10 min at room temperature or 100% methanol at -20° C for 20 min. Blocking, permeablization, antibody incubations and washing were done using blocking solution (10% fetal calf serum, 0.1%saponin and 0.02% sodium azide in PBS). The primary antibody was diluted in the blocking solution and incubated for 1 h at room temperature. The coverslips were washed 3-4 times with the blocking solution and the secondary antibody conjugated to fluorochromes was also diluted in the blocking solution and incubated at room temperature for 1 h in the dark. The coverslips were washed and mounted onto glass slides with mounting media and stored at in dark at 4° C until imaging. The cells were imaged using a Zeiss LSM-5 PASCAL laser scanning confocal microscope.

Deglycosylation experiments

Cells were lysed using native lysis buffer (25 mM Tris-HCl pH 8, 300 mM NaCl, 1% triton X-100 and protease inhibitor mix) for 30 min on ice and spun down at 13,000 rpm for 5 min. The cleared lysate was used for deglycosylation experiments. Deglycosylation of samples was performed using peptide N-glycosidase F (PNGase F) and endoglycosidase H (Endo H) according to manufacturer's recommendation. For tunicamycin (1 µg/ml) treatment, it was added to cells at the time of transfection. Digestion with the enzymes was carried out according to the manufacturer's protocol for 3 h at 37° C. The digest was then run on SDS-PAGE followed by western blotting as described before.

Table 2.1 List of primer sequences

Mutation	Forward primer sequence	Reverse primer sequence
N179Q	5’- ACTGGCAAGCAGTACACAATGGAA-3’	5’-TAATGTCGTCCTGAATCCAATGGCATC-3’
N192Q	5’- CAACTTGGCCAGTGTACATTTCCCCATCTCCGACC-3’	5’- GAAAAGTTCATACCATTCCATTGTGTAGTTCTTGCC-3’
N227Q	5’- TGGAAGGAACAGGGGACATTAGTTCAAGTAGCAAC-3’	5’-GTGAACATCATCAATTCCCTCAAAAAAGCAGGCAGCGCC-3’
N252Q	5’- TGGGTGAAACAGGACCAGGAAACAGGAATTTATTATGAG-3’	5’-CTTTGCCATTTGGTTGAACATGTTTCC-3’
N304Q	5’-ATAGAAACCCAGTATACAAGAATATTTCTTTACAGTGG-3’	5’-GTTCTTGAACTCTGATCCAAATTCAGCCAACTTG-3’
N320Q	5’-TACCTGGGACAGGAAACATCTGTTTTTGGGCC-3’	5’-GGTAGGTTCTCCACTGTAAAGAAATATTC-3’
N330Q	5’-CCAACAGGACAGAAGACTCTTGGTTTAGCC-3’	5’-CCCAAAAACAGATGTTTCATTTCCCAG-3’
N401Q	5’-CCTATCAGACAGAAAACACTCTCTGGT-3’	5’-TAAAGGGATTTCTTCATATGTTATTTT-3’
D279N	5’-GATTCCTACAACTGTTCCAAATTTGTGTTAAGGACC-3’	5’-AAACCATGTCTCTGCCCCCTTTTCTGGGCTGGC-3’
N192S	5’-CAACTTGGCAGCTGTACATTTCCCCATCTC-3’	5’- GAAAAGTTCATACCATTCCATTGTGTAGTTCTTGCC -3’
Y392X	5’-GGATCCGAGCTCGGTACCAAGC-3’	5’-TGTTATTTTAATAAAAGGGAATTTCATAGG-3’

Underlined bases indicate changed amino acid codon.

RESULTS

N-glycosylation of CLN5 *in vivo*

The Human CLN5 protein consists of 407 amino acids with eight putative N-glycosylation sites located at Asn 179, 192, 227, 252, 304, 320, 330 and 401. To determine which of these eight N-glycosylation sites is/are utilized *in vivo*, we removed each of these potential sites in *CLN5* cDNA by substituting Asn codon with Gln codon within the consensus sequence N-X-S/T. The *CLN5* cDNA with single N-glycosylation sites mutated (N179Q, N192Q, N227Q, N252Q, N304Q, N320Q, N330Q and N401Q) was used for transient expression in HeLa cells. We also created a mutant with an additional N-glycosylation site at Asp 279 (D279N), which is the same as a patient mutation seen mainly in the European population (12). The wild type CLN5 migrated on gel as a species of molecular weight of ~55 kDa. All eight N-glycosylation mutants showed a mobility shift corresponding to ~2.5 kDa reduction in size compared to wild type, indicating all eight putative N-glycosylation sites are indeed utilized *in vivo* (Fig. 2.1 A). The D279N mutant, as observed by others, migrated ~2.5 kDa higher than wild type CLN5.

To confirm that the reduction in size of mutants was due to the abrogation of N-glycosylation sites and not because of changing the amino acid to Gln, we changed Asn192 to Ser (N192S). The N192S mutant, which is also a patient mutation (10), showed similar reduction in size as the other single mutants (Fig. 2.1 B). Thus, these findings confirmed that all the eight potential N-glycosylation sites in CLN5 are used *in vivo* and removing them results in a reduction of ~2.5 kDa in the single mutant. We then created a double mutant with two of the N-glycosylation sites of CLN5 mutated (N192Q + N330Q). The double mutant ran ~2.5 kDa

lower than the single mutant and ~5 kDa lower than wild type CLN5 (Fig. 2.1 C), indicating that the reduction in size was due to the removal of two N-glycosylation sites.

Subcellular localization of CLN5 N-glycosylation mutants

CLN5 has been reported to be a lysosomal luminal protein and has been shown to reach the lysosomes via mannose-6 phosphate dependent and independent pathways (16). Glycosylation is important for proper folding and/or trafficking of luminal and secreted proteins. To assess the role of N-glycosylation in CLN5 localization, we examined subcellular localization of the N-glycosylation mutants. After 2 h of cycloheximide (CHX) chase (50 µg/ml), HeLa cells expressing CLN5 wt and mutants were fixed and immunostained for CLN5 with α-Myc antibody. Confocal microscopy analysis showed that the N192Q and N227Q mutants colocalized partially with the lysosomal marker Lamp2, similarly to CLN5 wt (Fig. 2.2 A and B). In our system, we have only been able to partially colocalize CLN5 with lysosomal markers. To address this issue, we used rabbit antiserum against CLN5 (IRmII-4) (13) to label for transiently-expressed CLN5 along with mouse antibody against Lamp1. CLN5 clearly colocalized with Lamp1 (Fig. S1), albeit the rabbit antiserum stained with higher cellular background compared to the mouse α-Myc antibody. N179Q, N252Q, N304Q and N320Q did not colocalize with lysosomal markerLamp2 (data not shown), but instead colocalized with ER marker Calnexin (Fig. 2.2 D). N330Q, on the other hand, was observed in the ER as well as in lysosomes (Fig. 2.2 E). The N401Q mutant colocalized with the Golgi marker Grasp65 (Fig. 2.2 C). These findings suggest that some of the N-glycosylation sites (N179, 252, 304, 320 and 330) are crucial for the

folding of CLN5, as obliterating these N-glycosylation sites results in CLN5 protein being retained in the ER, although N330Q had a milder phenotype. N401 appears to be essential for trafficking from the Golgi to the endosome/lysosomes, whereas N192 and N227 seem to have lesser roles in the folding or trafficking of CLN5.

To further assess the localization of wild type CLN5 as well as mutants that partially localized to the lysosome from above (N192Q, N227Q and N330Q), we decided to co-express these N-glycosylation mutants along with a mutant for Rab5 (eGFP-Rab5A Q79L). The small GTPase Rab5 gives the identity to early endosomes and has been used as an early endosome marker. The mutant Rab5A Q79L has a defective GTPase activity, resulting in the fusion and formation of enlarged endosomes (23, 24). Therefore, proteins destined to reach the lysosomes accumulate within the enlarged endosomes (luminal proteins) or on the limiting membrane of the enlarged endosomes (transmembrane proteins). HeLa cells were co-transfected for 24 h with eGFP-Rab5A Q79L and N-glycosylation mutants followed by 2 h of CHX chase before fixation. The lysosomal mutants (N192Q and N227Q) and wild type CLN5 localized to the inside of enlarged endosomes (Fig. 2.3 A), indicating that they truly localize to the endosome/lysosome. The ER mutants (N179Q, N252Q, N304Q and N320Q) were not detected inside the enlarged endosomes, but instead they were observed outside the enlarged endosomes (Fig 2.3 C). The mutant N401Q colocalized with Golgi marker (Grasp65) very well and did not localize inside the enlarged endosomes (Fig. 2.3 B). The N330Q was seen to localize equally to either inside or outside the enlarged endosomes (Fig. 2.3 D). These results conclusively show that N192Q and N227Q mutants can reach the endosome/lysosome, as does wild type CLN5. However, majority of the other mutants are either retained in the ER (N179Q,

N252Q, N304Q, N320Q) or cannot move forward beyond the trans-Golgi network (N401Q). We noted that small populations of all mutants escaped from the ER and were seen in the enlarged endosomes.

Based on the immunofluorescence imaging data, we categorized the N-glycosylation mutants depending on their localization. Table 2.2 summarizes the imaging data according to the colocalization with Lamp2, Calnexin, Grasp65 and whether the N-glycosylation mutant CLN5 localizes inside or outside of the enlarged endosomes of the mutant eGFP-Rab5A Q79L. Thus the N-glycosylation mutants can be categorized based on their subcellular localization as either lysosomal, ER or Golgi.

Deglycosylation studies of CLN5

It has been reported that CLN5 is Endoglycosidase H (Endo H) sensitive (14) and digestion with Endo H results in a reduction of the CLN5 molecular weight. After Endo H treatment, all the single mutants (including D279N) and wild type CLN5 had the same gel mobility of ~35 kDa (Fig. 2.4 A). This further confirmed that the 2.5 kDa size difference in mutants was due to removal of N-glycosylation, because removal of oligosaccharide chains eliminated the size difference between wild type and the mutants. This also implicates that all N-glycans on CLN5 are Endo H sensitive, as Endo H digestion on the mutants did not reveal any resistant chains (Fig. 2.4 A). If there were any Endo H resistant chains, then N-glycosylation mutant(s) of that site(s) would have had a lower gel mobility compared to CLN5 wt band. Complex and some hybrid sugars are resistant to Endo H digestion, suggesting that the CLN5 N-glycosylation sites mostly consist of

high mannose type sugars. The difference in the size of wild type CLN5 and Endo H digested CLN5 was ~20 kDa (Fig. 2.1 A and 2.4 A), which can be explained by eight oligosaccharide chains of ~2.5 kDa each observed in CLN5.

To investigate N-glycosylation modification further, we digested CLN5 with peptide N-glycosidase F (PNGase F). Endo H cuts after the first N-acetyl glucosamine of the oligosaccharide chain on the Asn, thus leaving one sugar moiety (GlcNAc) on Asn. In contrast, PNGase F cuts immediately after the Asn and does not leave any sugar moiety. When we treated wild type CLN5 with PNGase F, we observed faster migration of CLN5 compared to Endo H treatment, consistent with eight N-acetyl glucosamine moieties being further removed (Fig. 2.4 B). We also examined the effect of tunicamycin treatment (1 μg/ml) on CLN5 expression. Tunicamycin prevents the first step of N-glycosylation in the ER, thus preventing any N-glycosylation from taking place. Interestingly, we observed a faster migration of tunicamycin treated CLN5 wt (~32 kDa) compared to PNGase F treated CLN5 wt (Fig 2.4 B). This result indicates that there is/are other modification(s) on CLN5 beyond the N-glycosylation. It should be noted that we cannot rule out the possibility of the size difference being due to PNGase F digestion converting the Asparagine to Aspartic acid on each N-glycosylation site.

Since tunicamycin treatment caused CLN5 being retained in the ER (data not shown), any modification beyond the ER would not occur. To test if this extra modification happens after CLN5 exits the ER, we used N-glycosylation mutants that primarily localized to the ER so as to examine their modification. Surprisingly, when we digested N320Q with PNGase F, we observed faster mobility shift compared to other ER mutants and CLN5 wt (Fig. 2.4 C), although

tunicamycin treated sample still ran lower than N320Q. This indicates that there is a possibility of a PNGase F resistant residue being present on Asn320 which is abolished in the N320Q mutant resulting in the migration difference, although a PNGase F resistant linkage is rare in animals and common in plants. These experiments show that there are potentially other modifications apart from N-glycosylation that could be taking place on CLN5 and that these modifications possibly take place post ER or post N-glycosylation.

Stability of CLN5 deficient in N-glycosylation

In immunofluorescence studies, we noted that the lysosomal mutants had lower CLN5 signal than ER and Golgi mutants after 2 h of CHX chase. To investigate if the lysosomal mutants were being degraded or turned over more efficiently than the other mutants, transient transfection of HeLa cells was performed for 24 h followed by CHX chase for 0, 2 and 4 h. Since we and others have previously detected CLN5 in media (secreted out), we changed media at time of CHX chase and collected media at 2 and 4 h to examine if lower protein levels correspond to an increase in secreted CLN5. The lysosomal mutant protein levels gradually decreased in the pellet with increasing time of CHX chase and this corresponded to the increase in CLN5 being detected in the media with increasing time of CHX treatment (Fig. 2.5). Among the ER mutants, N179Q, N252Q, and N330Q showed stable protein levels in the pellets at all time points without any protein being detected in the media. While N304Q, N320Q, and D279N patient mutant showed some degree of reduced protein level in the pellet, interestingly, for N401Q the decreasing levels in the pellets and the increasing levels in the media were more dramatic than the lysosomal mutants (N192Q and N227Q) and CLN5 wt. These results show

that N-glycosylation has varied effect on the stability of CLN5, indicating that some of the N-glycosylation sites are essential for the proper folding and stability of CLN5 and depleting those sites results in CLN5 being degraded.

Subcellular localization of CLN5 patient mutations

Since N192Q mutant localizes to the lysosome just like wild type CLN5 does, we decided to examine the localization of patient mutation N192S. Compared to other patient mutations D279N and Y392X that localized to the ER, the N192S was observed in the enlarged endosomes in cells expressing eGFP-Rab5A Q79L (Fig. 2.6). This suggests that, unlike D279N and Y392X that cannot fold properly and therefore are retained in the ER, the N192S mutant has functional defect in the lysosome. This was the first patient mutation that was characterized and shown to localize to the lysosome. The D279N mutant localized outside of the enlarged endosomes and also with ER marker Calnexin (Fig. 2.6). Y392X mutant did not localize to inside of enlarged endosomes nor with Golgi marker Grasp65 (Fig. 2.6).

DISCUSSION

While the function of CLN5 is not known at the moment, it is clear that several modifications occur during or after translation which might have an effect on its function, stability and trafficking. Based on experimental evidence and sequence prediction, it is clear that the N-terminus undergoes signal peptide cleavage in the ER during co-translation (16).

From the available sequences, it is apparent that signal peptide of CLN5 in primate is unusually long (see alignment in supplementary figure S2) compared to other species. Another major modification is N-glycosylation. In this report, we demonstrate that all eight putative N-glycosylation sites of human CLN5 are utilized *in vivo*. Seven of the eight N-glycosylation sites are conserved among mammalian species (supplementary Fig. S2). In *Mus musculus* and *Rattus norvegicus*, the N-glycosylation site corresponding to human N401 is not present. We have uncovered that N401 is essential for lysosomal trafficking of human CLN5, as N401Q predominantly localized to the Golgi after 2 hour CHX chase (Fig. 2.2 C and 2.3 B) and majority of N401Q CLN5 was secreted out to the media (Fig. 2.5). Interestingly, mouse CLN5 has been shown to reach lysosomes in M6PR independent manner (16). Asparagines on residue 401, 320, and 330 of human CLN5 were predicted to have a M6P tag attached in a proteomic scaled mass spectrometry analysis (21). However, from our results we know that N330Q can reach beyond the endosome (Fig. 2.2 E and 2.3 D), indicating that this residue may not be involved in M6PR-mediated transport. The N-glycosylation on N320 is essential for proper folding, as N320Q was retained in the ER (Fig. 2.2 D and 2.3 C). Therefore we cannot evaluate its involvement in TGN to endosome trafficking. N401, on the other hand, has a clear role in transport from Golgi to the endosome as N401Q mutant was either retained in the Golgi (Fig. 2.2 C and 2.3 B) or secreted out of the cell (Fig. 2.5), similar to the effect of I-cell disease where loss of M6P tag leads to most lysosomal acid hydrolases being secreted out of the cell. N-glycosylation on residue 401 also confirms that CLN5 is not a transmembrane protein as some prediction program suggested, because a single transmembrane domain at region 352-373 would position N401 at the cytoplasmic side.

CLN5 has been reported to localize to the lysosome (13, 14). However, in our system, we can only partially co-localize CLN5 with lysosomal markers. In attempt to address this issue, we used rabbit antiserum against CLN5 (IRmII-4) (13) to label overexpressed CLN5 (this antiserum, as well as other CLN5 antibodies, do not recognize endogenous CLN5). There were punctate structures labeled with CLN5 that clearly colocalized with lamp1 (Fig. S1). CLN5 also partially colocalizes with endosome markers EEA1 and retromer (data not shown). With eGFP-Rab5A Q79L co-expression, we have shown CLN5 to localize to the inside of enlarged endosomes, verifying its presence in the endosome/lysosome pathway(s). A recent study indicated a role of CLN5 in regulating retromer recruitment to the endosomes (15). When CLN5 was depleted, the cellular levels of endosomal/lysosomal cargo receptors: cation-independent M6PR (CI-M6PR) and sortilin were reduced. Our colocalization data is consistent with a function of CLN5 in the endosomes.

Using Endo H glycosidase, we showed that all eight N-glycans are Endo H sensitive, because by deleting individual N-glycosylation, we did not observe lower molecular weight product when mutant CLN5 was treated with Endo H. This can be explained by the fact that if there is an Endo H resistant N-glycan and when that specific N-glycosylation site is removed, we should observe a lower molecular weight product compared to wild type CLN5 upon Endo H digestion. However, with Endo H digestion, all mutants and wild type CLN5 ran the same without any mobility shift. Intriguingly, a PNGase F resistant N-glycan was identified on N320, as digestion with PNGase F of N320Q mutant resulted in a lower molecular weight species. PNGase F cannot digest the core N-acetyl-glucosamine that is modified by an α 1-3 fucose. However this modification is rare in animal cells. Another interesting finding is that with tunicamycin

treatment, we observed an even lower molecular weight CLN5. Tunicamycin inhibits the enzyme GlcNAc phosphotransferase and therefore prevents N-glycosylation. In tunicamycin-treated cells, CLN5 localizes to the ER (data not shown). It is possible that due to retention in the ER, CLN5 is not accessible to other modifications that occur past the ER which would be consistent with tunicamycin treated CLN5 wt running lower that PNGase F treated CLN5 wt.

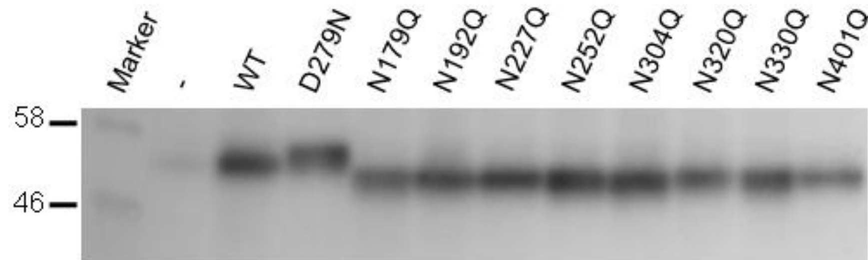
Previous characterized patient mutations with aberrant N-glycosylation include D279N, N192S, and Y392X. We show that D279N mutation introduces an extra N-glycosylation site to CLN5 which causes an increase in its molecular weight (Fig. 2.1 A). This causes CLN5 to be retained in the ER (Fig 2.6). Y392X has been reported to localize to the ER as well as the Golgi (12). Since this slightly truncated version of CLN5 does not have the N401 residue, it's possible that its main defect is the missing N-glycan on N401 residue and it would behave similar to the N401Q mutant. However we observed ER retention of the Y392X mutant (Fig. 2.6), indicating that missing the last 15 amino acids of CLN5 has a more dramatic effect on the overall folding compared to the N401Q mutant. In this report we also characterized the localization of a patient mutation N192S. Similar to N192Q N-glycosylation mutant, we observed normal lysosomal/endosomal localization of N192S (Fig. 2.6). This suggests that N192S is a mutant with defect in its lysosomal function, rather than general folding issues as seen with other patient mutations of CLN5. This opens up new avenues to explore the function and role of CLN5 in the lysosome using N192S mutant as a tool for these studies.

This study explores the role of N-glycosylation and its importance in the stability and transport of CLN5 protein. It aids the field of NCL and CLN5 to move forward by characterizing

some of the basic facets of human CLN5 protein. Although the function of CLN5 is not known, it is possible that N-glycosylation plays a major role in the proper functioning of CLN5 in the lysosomes, similar to TPP1 (18).

Figure 2.1

A



B

C

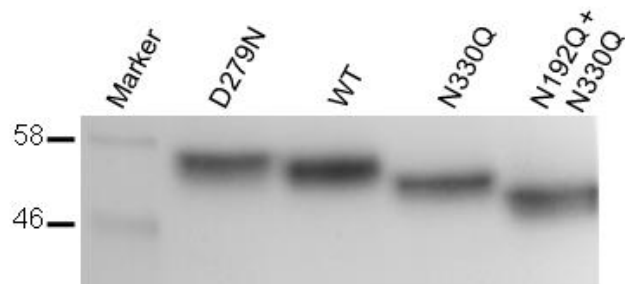
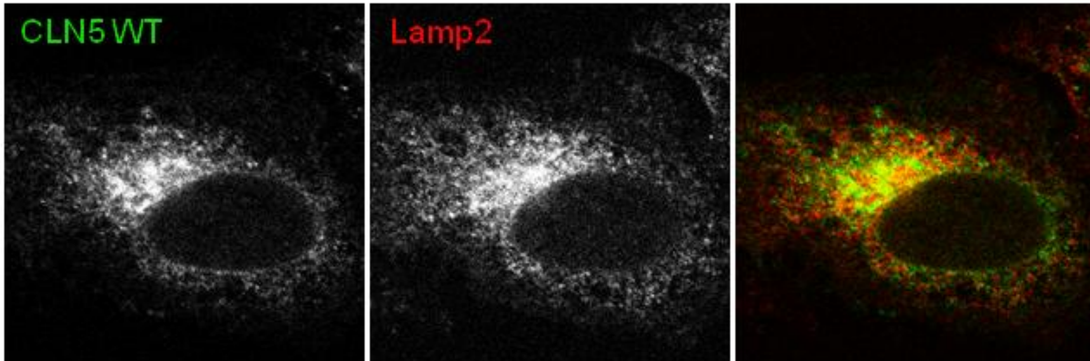


Figure 2.1- N-glycosylation of CLN5 *in vivo*. (A) Western blot analysis of CLN5 wt and N-glycosylation mutants. HeLa cells were transiently transfected for 24 h and the whole cell lysate (WCL) was collected used for western blotting analysis using mouse α-Myc antibody to detect CLN5. Equal amount of protein was loaded onto each well. (B) Western for confirming mobility shift is due to abrogation of the N-glycosylation site and not due to changing of amino acid to Gln. HeLa cells were transiently transfected with CLN5 wt, N192Q and N192S (patient mutation) for 24 h and the WCL was run on gel for western to detect mobility shift. (C) Western of double mutant to show gradual size reduction of CLN5. HeLa cells were transfected for 24 h with CLN5 wt, D279N, N330Q and N192Q +N330Q (double mutant) and the WCL was resolved on a gel for western. These data are representative of at least two independent experiments.

Figure 2.2

A



B

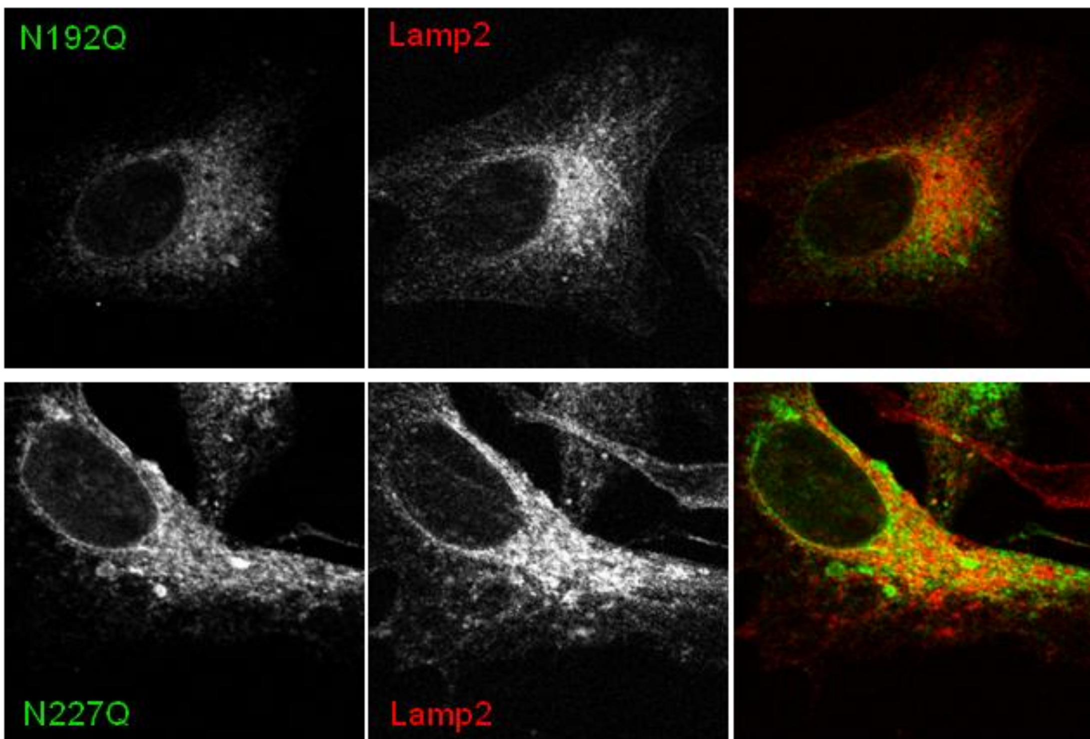


Figure 2.2

C

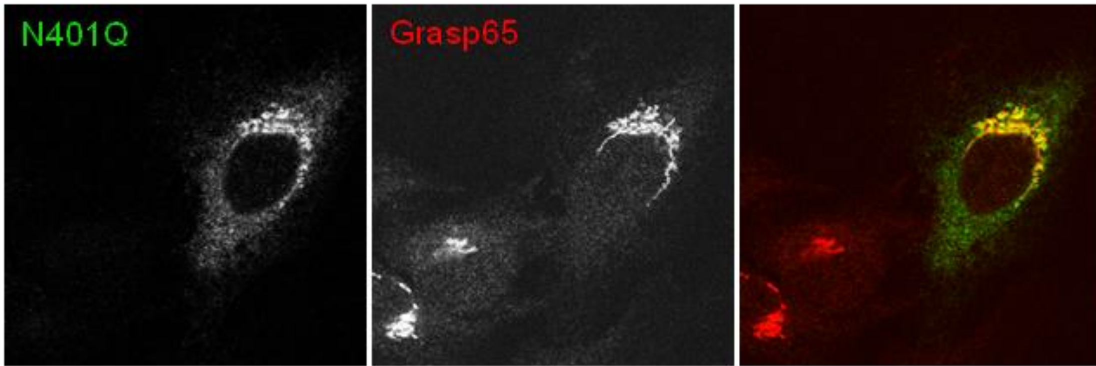


Figure 2.2

D

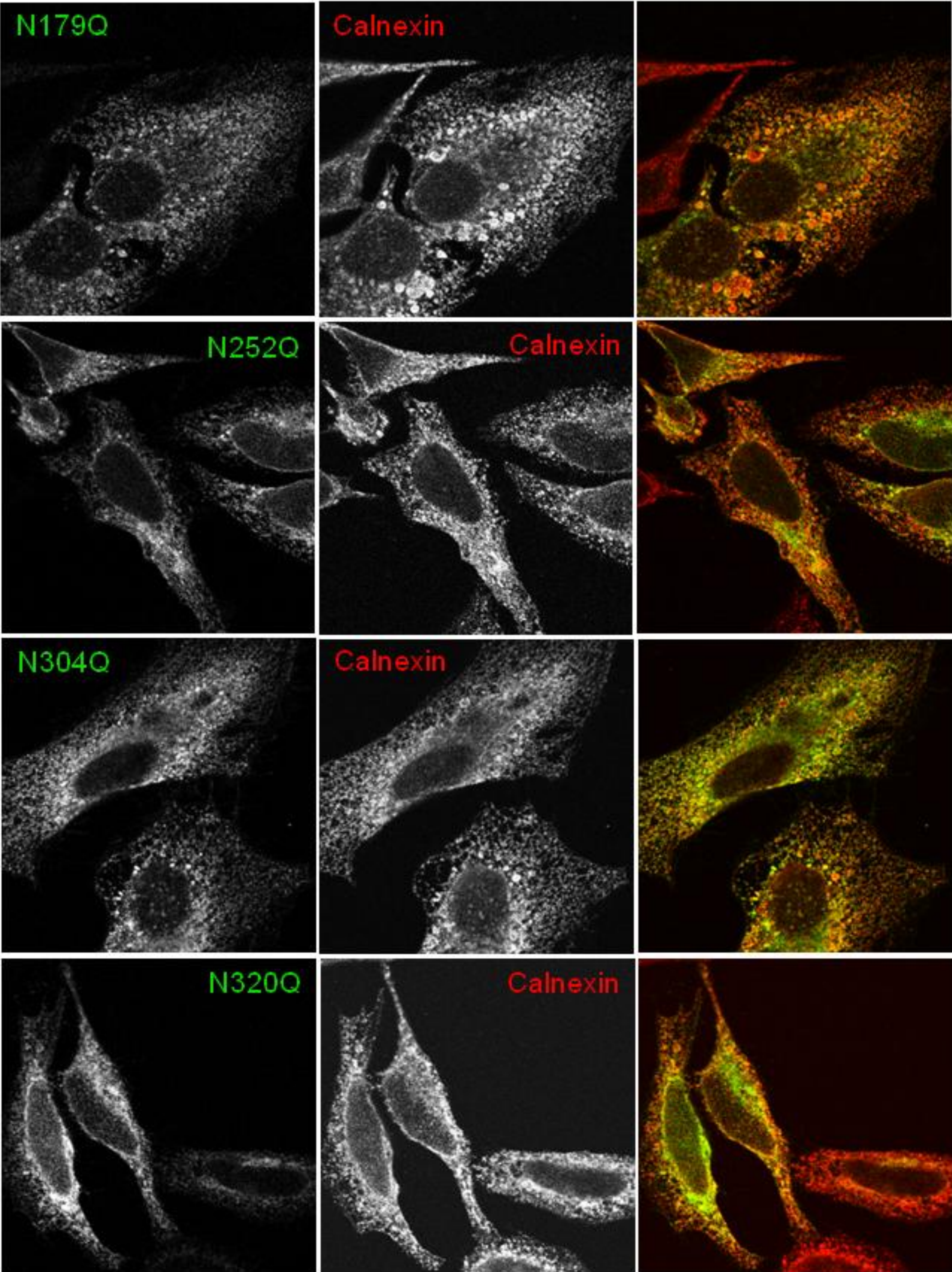


Figure2.2

E

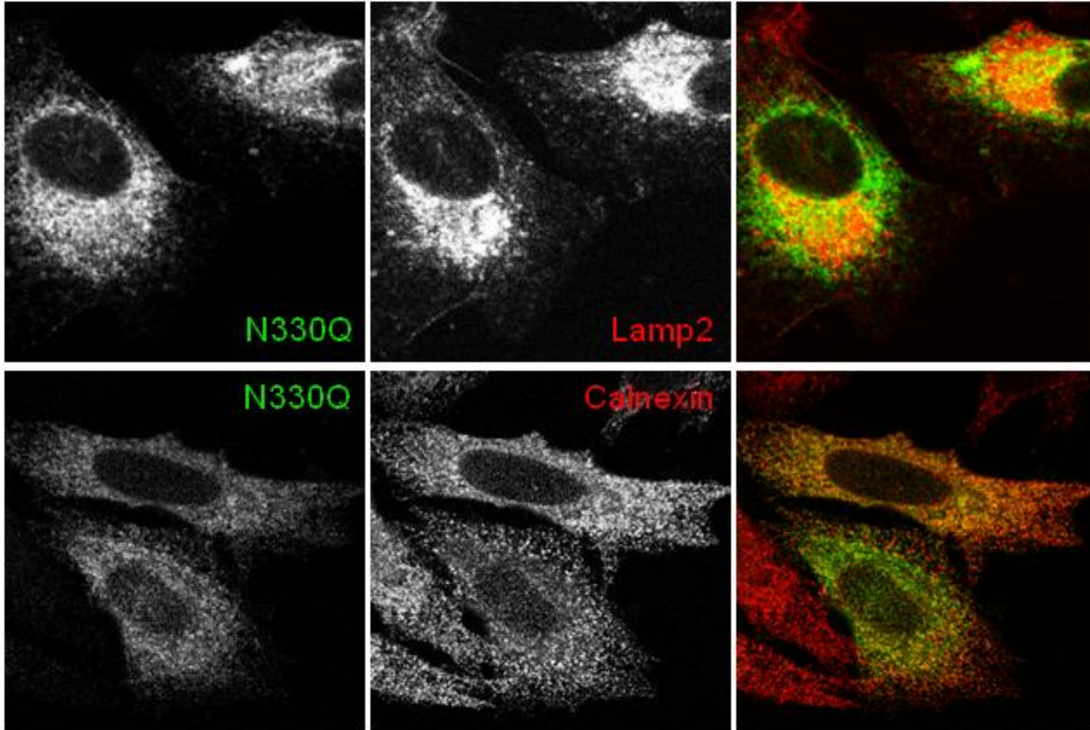


Figure 2.2- Subcellular localization of CLN5 N-glycosylation mutants. Immunofluorescence microscopy of transiently transfected N-glycosylation mutant HeLa cells. HeLa cells were seeded on glass coverslips and transfected with CLN5 wt or mutants for 24 h. The cells were treated with cycloheximide (50 µg/ml) for 2 h and fixed using 4% formaldehyde and stained for CLN5 (mouse α-Myc antibody) , ER marker Calnexin, Lysosome marker Lamp2 or Golgi marker Grasp65. The cells were imaged on a Zeiss LSM-5 PASCAL laser scanning confocal microscope. (A) CLN5 wt colocalizes partially with lysosomal marker Lamp2 (yellow). N192Q and N227Q mutants also partially colocalized with lamp2 (B). (C) N401Q colocalizes with the Golgi marker Grasp65 while the other N-glycosylation mutants – N179Q, N252Q, N304Q, and N320Q colocalized with ER marker Calnexin (D). N330Q colocalized partially with both Lamp2 and Calnexin (E). These data are representative of at least two independent experiments (1000X magnification).

Figure 2.3

A

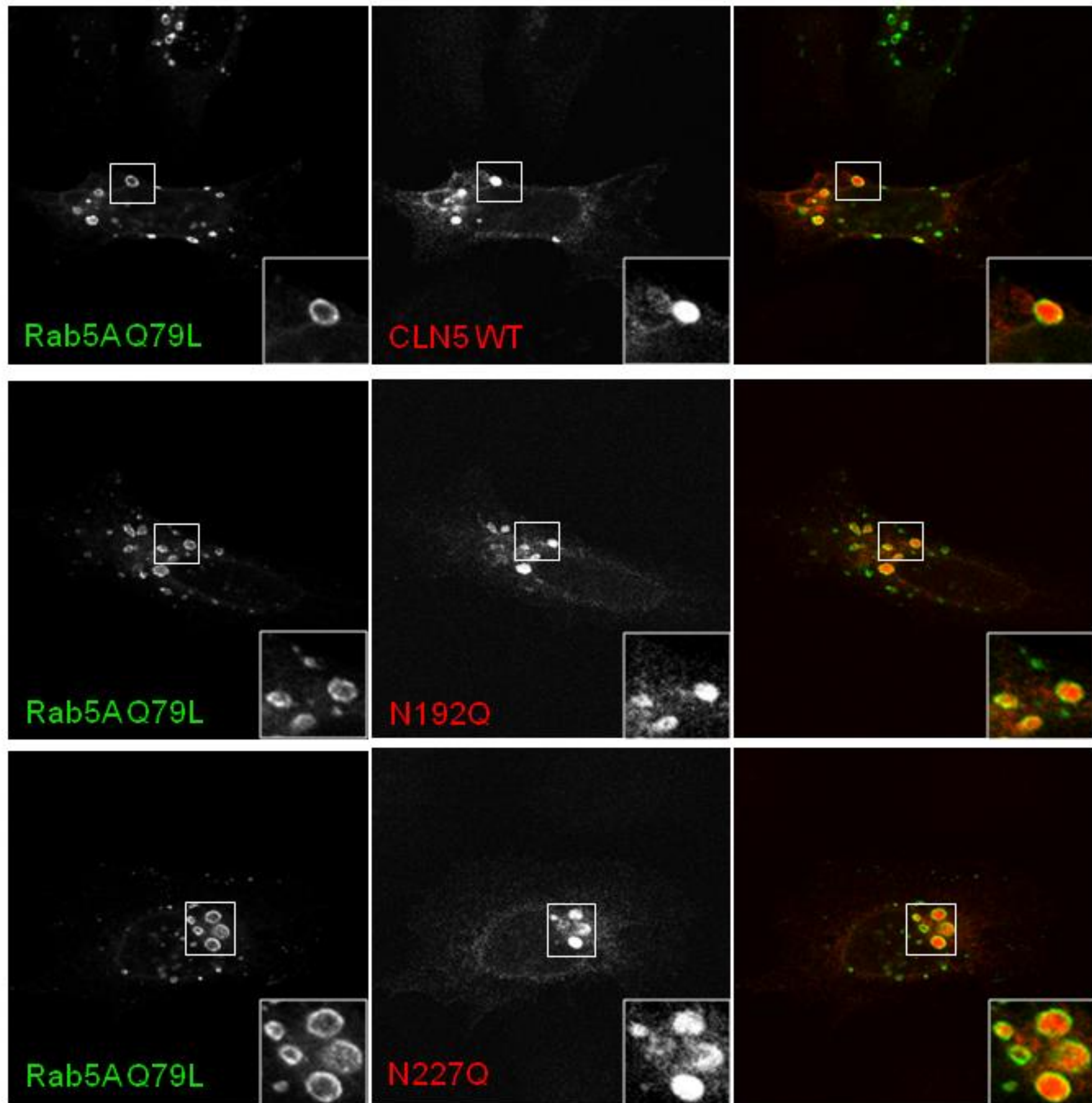


Figure 2.3

B

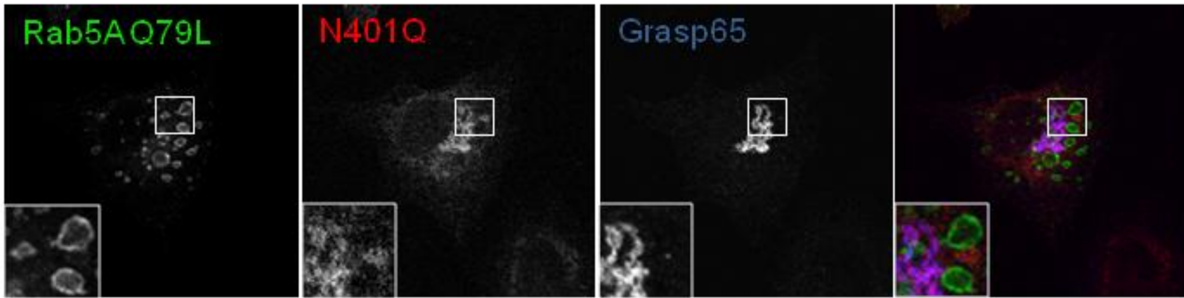


Figure 2.3

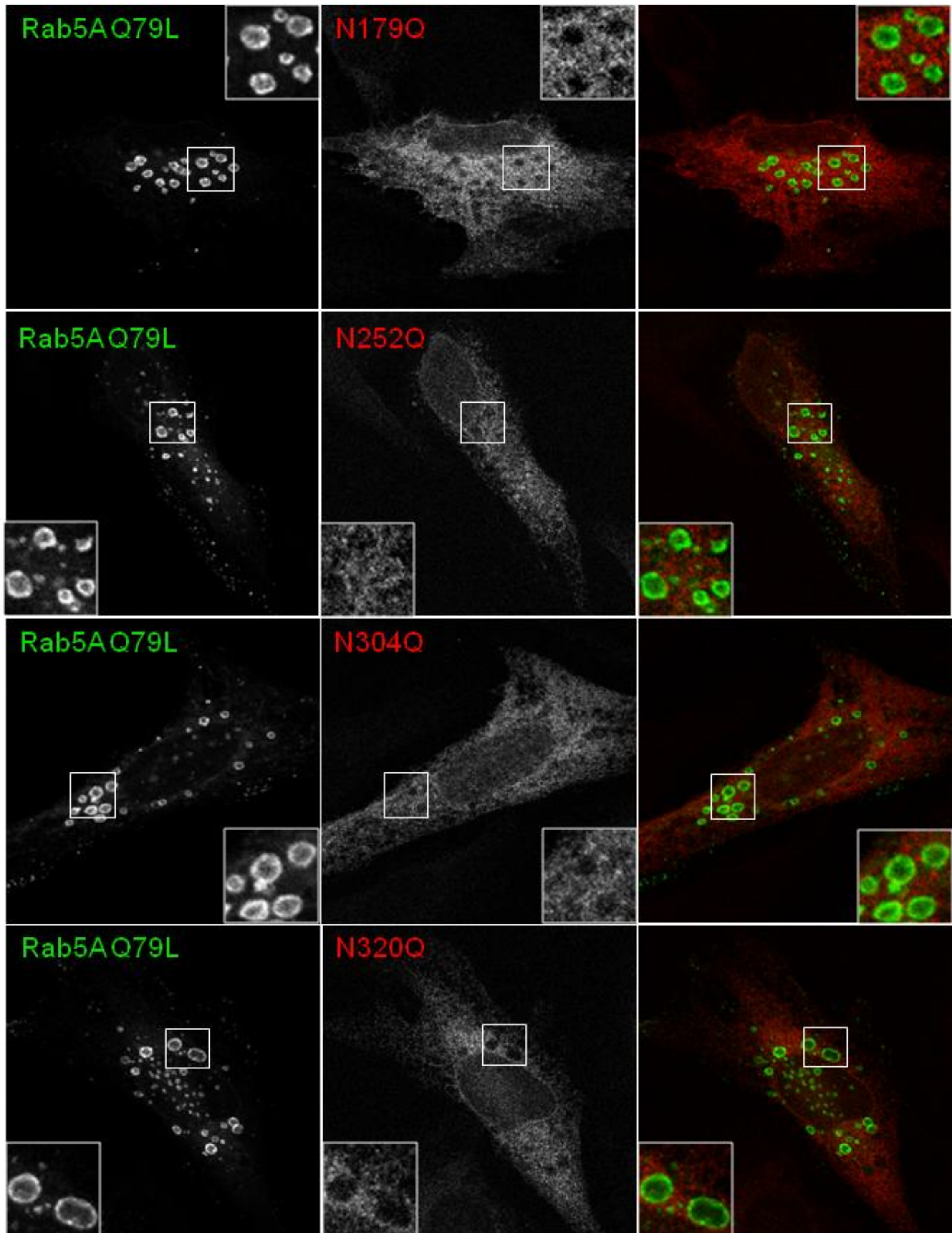


Figure 2.3

D

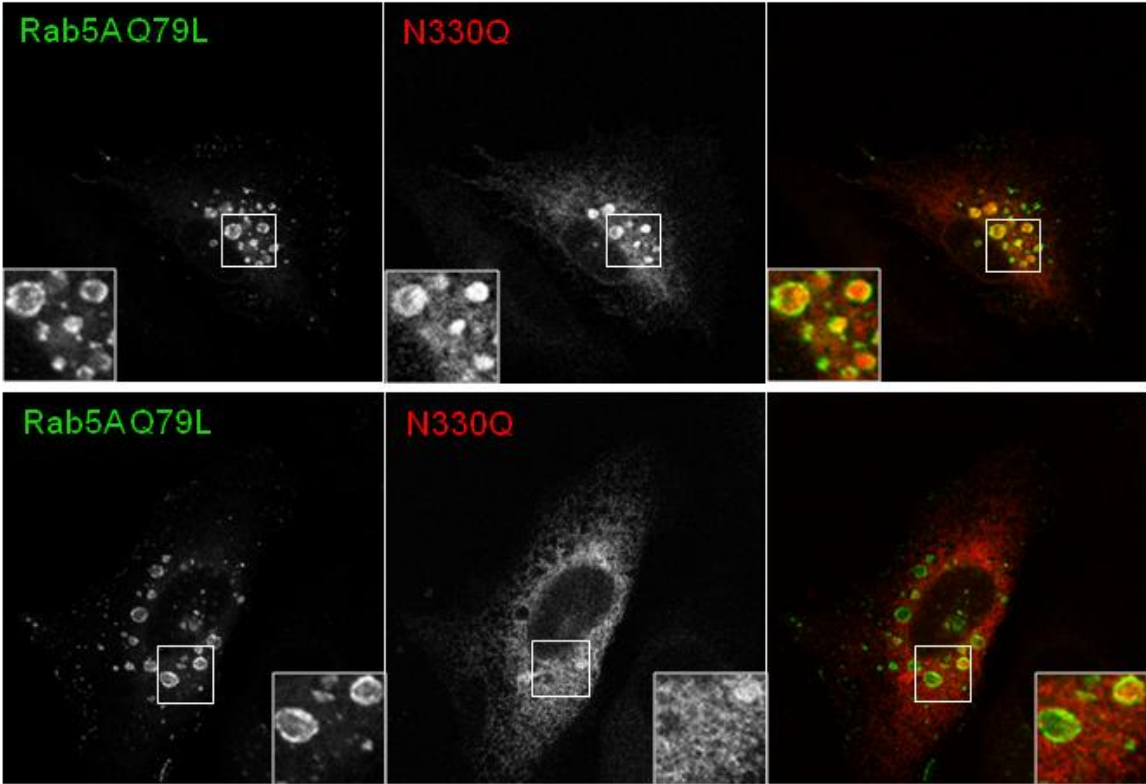


Figure 2.3- Accumulation of lysosomal-localized CLN5 in rab5A Q79L-induced enlarged endosomes. We used a Rab5A mutant (eGFP-Rab5A Q79L) that is defective in its GTPase activity and leads to the formation of enlarged endosomes. HeLa cells were seeded on glass coverslips and double transfected with eGFP-Rab5A Q79L and CLN5 wt or N-glycosylation mutants for 24 h. The cells were chased with CHX for 2 h and fixed using 4% formaldehyde. The cells were then stained for CLN5 (mouse α-Myc antibody) and imaged on a Zeiss LSM-5 PASCAL laser scanning confocal microscope. (A) CLN5 wt, N192Q and N227Q can be seen to localize within the enlarged endosomes. N401Q did not localize with the enlarged endosomes, instead colocalizing with Golgi marker Grasp65 (C). N179Q, N252Q, N304Q, and N320Q were not seen to localize within the enlarged endosomes indicating that they are stuck in the ER (D). N330Q was seen to localize insides and outside the enlarged endosomes (E), consistent with our previous results. These data are representative of at least two independent experiments (1000X magnification).

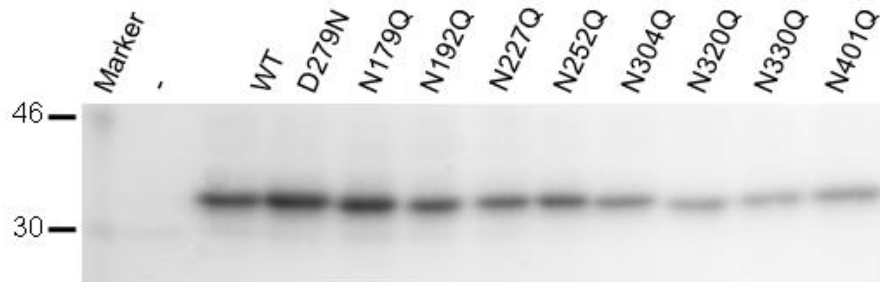
Table 2.2

Mutant/ WT	Lamp2 (Lysosome)	Calnexin (ER)	Grasp65 (Golgi)	Inside or Outside of enlarged endosomes	Localization
CLN5 WT	+/-	-	-	Inside	Lysosome
N179Q	-	+ +	-	Outside	ER
N192Q	+/-	-	-	Inside	Lysosome
N227Q	+	+/-	-	Inside	Lysosome
N252Q	-	+ +	-	Outside	ER
N304Q	-	+ +	-	Outside	ER
N320Q	-	+ +	-	Outside	ER
N330Q	+/-	+	-	Inside/Outside	Lysosome/ER
N401Q	-	-	+ +	Outside	Golgi

Table 2.2- Summary of subcellular localization of CLN5 N-glycosylation mutants. Based on the imaging data from figure 2.2 and whether CLN5 localizes inside or outside of enlarged endosomes (Figure 2.3), N-glycosylation mutants have been categorized into either being in the lysosome, ER or Golgi. + indicated positive colocalization, - indicates no colocalization, +/- indicates partial colocalization.

Figure 2.4

A



B

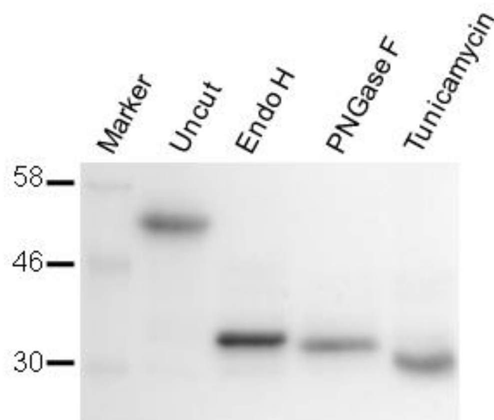


Figure 2.4

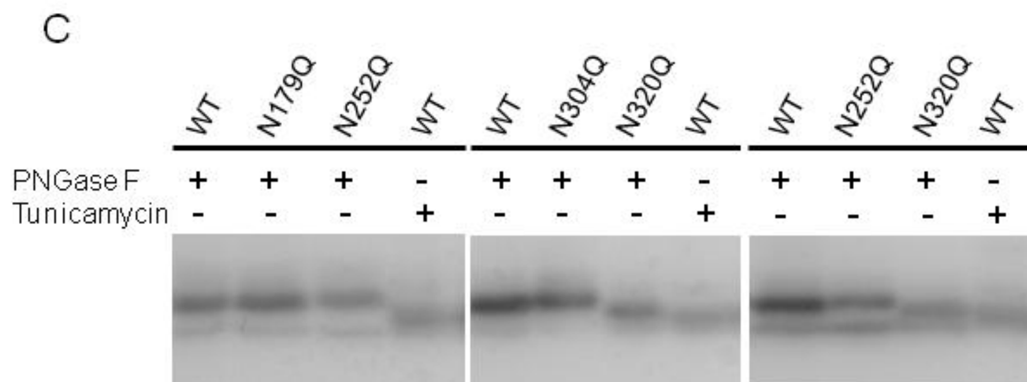


Figure 2.4- Endo H sensitivity and deglycosylation studies of CLN5. Transiently transfected (24 h) HeLa cells were used for these studies. HeLa cells were transfected with CLN5 wt or N-glycosylation mutants for 24 h and the WCL was used to treat with Endo H enzyme (A) for 3 h at 37° C after which it was run on gel and detected by western blotting using mouse α-Myc antibody. (B) CLN5 wt transfected HeLa WCL was used for cutting with Endo H, PNGase F for 3 h at 37° C along with tunicamycin (1 µg/ml) treated CLN5 transfected WCL. Tunicamycin was added to cells at the time of transfection. (C) ER mutants N179Q, N252Q, N320Q, and N304Q along with CLN5 wt were used to transfect HeLa cells for 24 h and the WCL was used to cut with PNGase F. CLN5 wt tunicamycin sample was also used. These data are representative of at least two independent experiments.

Figure 2.5

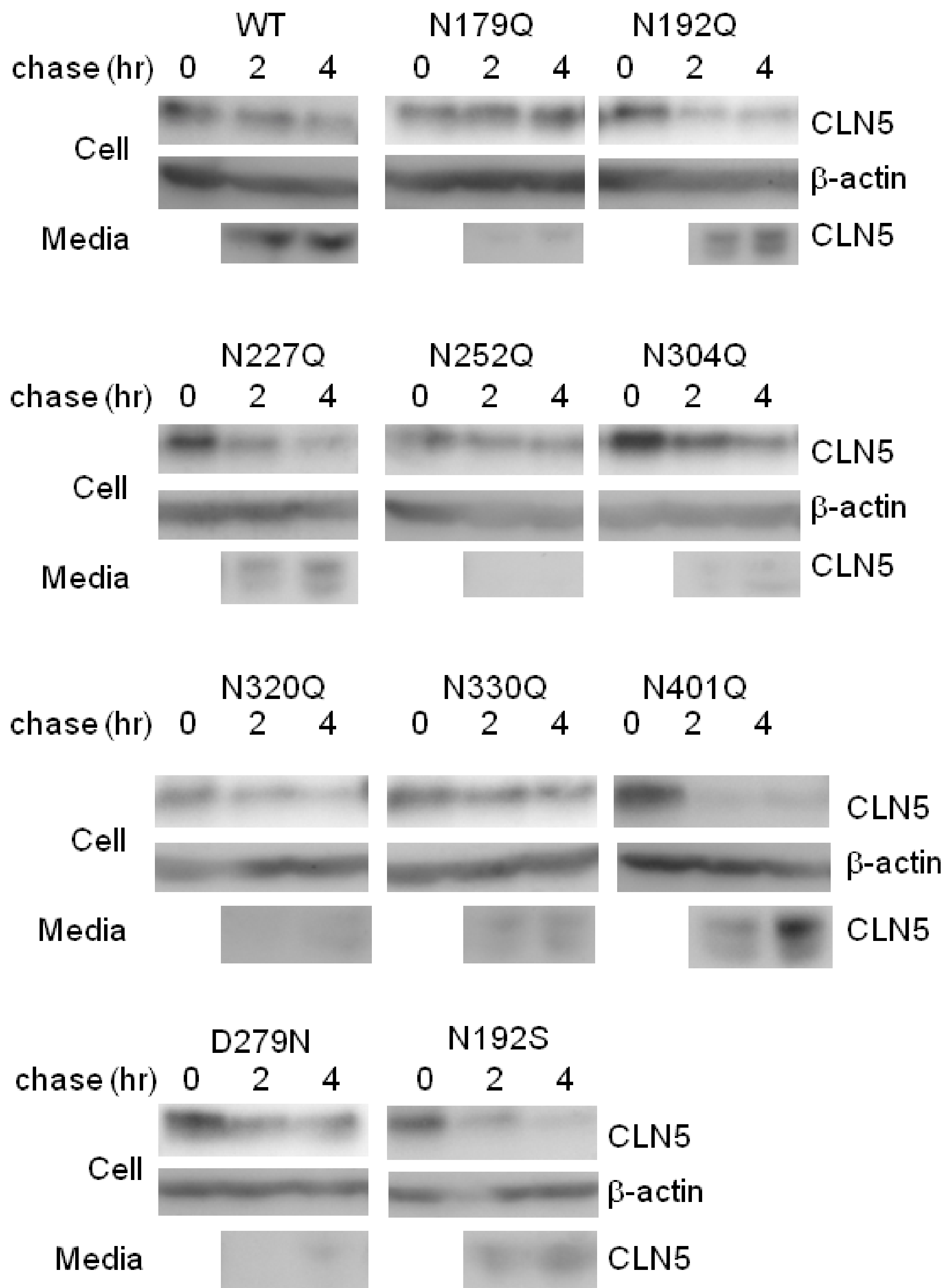


Figure 2.5- Stability of N-glycosylation deficient CLN5 protein. HeLa cells were transfected with CLN5 wt or N-glycosylation mutants for 24 h after which CHX chase was performed for 0, 2 , and 4 h. The media was changed at the time of CHX chase. Media was collected at 2 and 4 h of CHX chase and the pellet was collected at 0, 2, and 4 h time points. The media samples were precipitated using ammonium sulfate and run on gel for western. The WCL of pellets was also run on gel for western and CLN5 was detected using mouse α-Myc antibody. β-actin was used as loading control for the pellet samples. These data are representative of at least two independent experiments.

Figure 2.6

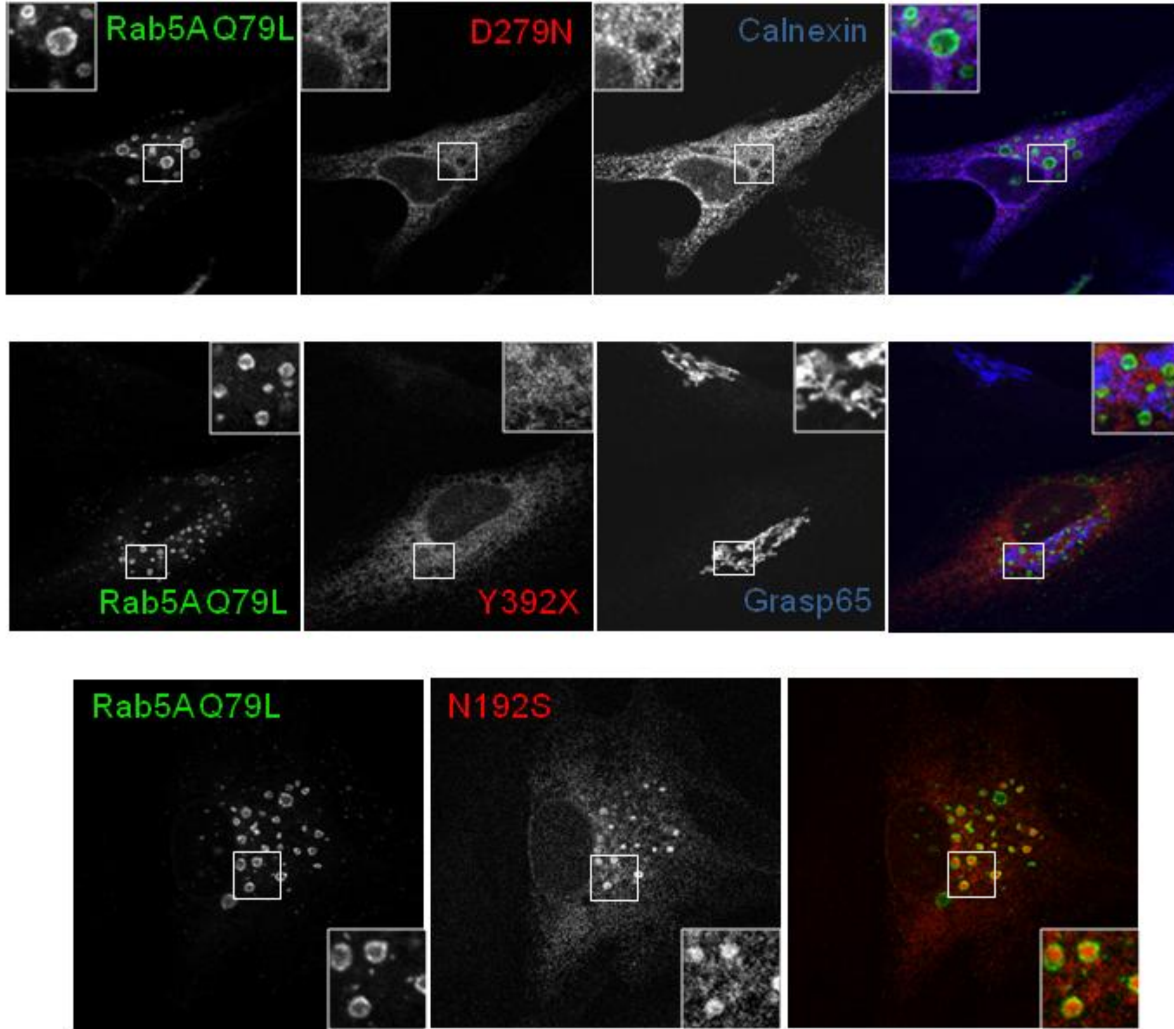
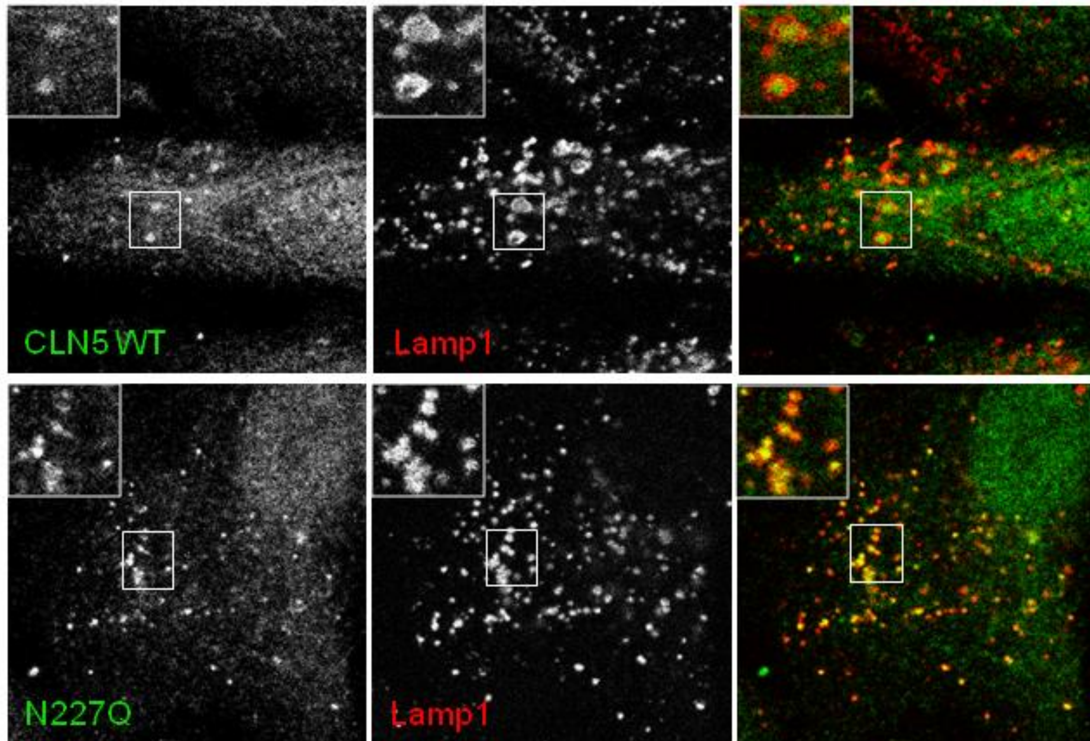


Figure 2.6- Subcellular localization of CLN5 patient mutations. HeLa cells were seeded on glass coverslips and double transfected using eGFP-Rab5A Q79L and D279N, Y392X or N192S for 24 h after which they were chased with CHX for 2 h and fixed using 4% formaldehyde and stained for CLN5 using mouse α-Myc antibody. They were then imaged on a Zeiss LSM-5 PASCAL laser scanning confocal microscope. D279N did not localize inside of enlarged endosomes and was similar to the ER mutants. Y392X also did not localize inside of the enlarged endosomes whereas N192S mutant localized to the inside of enlarged endosomes similar to CLN5 wt, N192Q, and N227Q mutants (1000X magnification).

Figure S1



Supplementary Figure S1- Subcellular localization of CLN5 wt. HeLa cells were seeded on glass coverslips and transfected with CLN5 wt for 24 h. The cells were then chased with CHX for 2 h and fixed with 100% methanol for 20 min at -20° C. They were then stained for CLN5 using rabbit α-CLN5 Ab (IRmII-4) and mouse α-Lamp1. The cells were imaged on a Zeiss LSM-5 PASCAL laser scanning confocal microscope (1000X magnification).

Supplementary Figure S2

```
CLUSTAL multiple sequence alignment

H. sapien       ------------------------------------------------------------
P. troglodytes  ------------------------------------------------------------
M. mulatta      MVEESEKRSSVLGGAPVTDFRFWFGFGGGCWAGTGQFIKEEAPGLDRILLGARPKRLPGL 60
B. taurus       ------------------------------------------------------------
O. aries        ------------------------------------------------------------
C. familiaris   ------------------------------------------------------------
C. jacchus      MRALGTLREETT-EARPAVLTTDQGQPEFVFVALGSPELNGKFCCIRFHYLSQTKAPSWT 59
M. musculus     ------------------------------------------------------------
R. norvegicus   ------------------------------------------------------------

H. sapien       --------------MRRNLRLGPSSGADAQGQGAPRPGLAAPRMLLPPASQASRGSGSTG 46
P. troglodytes  --------------MRRNLRLLPSSGADAQGQGAPRPGLAAPRMLLPPASQASRGSGSTG 46
M. mulatta      KSAAAAAGRVRRGVTRRNLRLLPSSGPDXPGWGVPRPGLAAPRMLLPQASQASRGSGSTG 120
B. taurus       ------------------------------------------------------------
O. aries        ------------------------------------------------------------
C. familiaris   ------------------------------------------------------------
C. jacchus      AKCGSRRGQGAGRSHALEHGSLPNSGASA-GHGGLRPGLAAPRKLHPPASRGAGGAGSTG 118
R. norvegicus   ------------------------------------------------------------

H. sapien       CSLMAQEVDTAQGAEMRRGAGA--ARGRASWCWALALLWLAVVPGWSRVSGIPSRRHWPV 104
P. troglodytes  CSLMAQEVDTAQGAEMRRGAGA--ARGRASWCWALALLWLAAVQGWSRASGIPSRRHWPV 104
M. mulatta      CSLMAPELDTAQGAKMRRGAGA--ARGRASWCPAVALLWLAAVPGWPRASGVLSRRHWPV 178
B. taurus       ---MAQVGSAGPGACGRRGAGAGAGPERTTWRWAPALLWLATA---AAVAGDPSRRQWPV 54
O. aries        ---MAQAGGAGAGAWGRRGAGAGAGPERAPWRWAPALLWLAAATAAAAAAGDPSRRQWPV 57
C. familiaris   ---MAQAGSADPGVGGHWAAG----PRCAPWRWALALLWLATA------AGGPSRRQWPV 47
C. jacchus      GSRMAQEGDTALGVEMRGGAGG--ARGLASRRWALALLFLAAAPGWLPASGSPSRRHWPE 176
M. musculus     --------------MLRGGPCG------AHWRPALALALLGLAT--ILGASPTSGQRWPV 38
R. norvegicus   --------------MPRVGPGG------AHWRPALALALLGLAA--TLGASPTSGQRWPV 38
                                : ..        :    * **  *. .      :.  * ::** 

H. sapien       PYKRFDFRPKPDPYCQAKYTFCPTGSPIPVMEGDDDIEVFRLQAPVWEFKYGDLLGHLKI 164
P. troglodytes  PYKRFDFRPKPDPYCQAKYTFCPTGSPIPVMEGDDDIEVFRLQAPVWEFKYGDLLEHLKI 164
M. mulatta      PYKRFDSRPKPDPYCQAKYTFCPTGSPIPVMEGDDDIEVFRLQAPVWEFKYGDLLGHLKI 238
B. taurus       PYKRFSFRPEPDPYCQAKYTFCPTGSPIPVMKDDDVIEVFRLQAPVWEFKYGDLLGHLKI 114
O. aries        PYKRFSFRPEPDPYCQAKYTFCPTGSPIPVMKDDDVIEVFRLQAPVWEFKYGDLLGHLKI 117
C. familiaris   PYKRFSFRPEPDPYCQAKYTFCPTGSPIPVMKGDDVIEVFRLQTPVWEFKYGNLLGHLKI 107
C. jacchus      PYKRFDFRPKPDPYCQAKYTFCPTGSPIPVMKDDDDIEVFRLQAPVWEFKYGDLLGHLKI 236
M. musculus     PYKRFSFRPKTDPYCQAKYTFCPTGSPIPVMKDNDVIEVLRLQAPIWEFKYGDLLGHFKL 98
R. norvegicus   PYKRFSFRPKTDPFCQAKYTFCPTGSPIPVMKDNDVIEVLRLQAPVWEFKYGDLLGHFKI 98
                *****. **:.**:*****************:.:* ***:***:*:******:** *:*:

H. sapien       MHDAIGFRSTLTGKNYTMEWYELFQLGNCTFPHLRPEMDAPFWCNQGAACFFEGIDDVHW 224
P. troglodytes  MHDAIGFRSTLTGKNYTMEWYELFQLGNCTFPHLRPEMDAPFWCNQGAACFFEGIDDVHW 224
M. mulatta      MHDAIGFRSTLTGKNYTMEWYELFQLGNCTFPHLRPEMDAPFWCNQGAACFFEGIDDVHW 298
B. taurus       MHDAIGFRSTLTEKNYTMEWYELFQLGNCTFPHLRPEMNAPFWCNQGAACFFEGIDDSHW 174
O. aries        MHDAIGFRSTLTEKNYTMEWYELFQLGNCTFPHLRPEMNAPFWCNQGAACFFEGIDDNHW 177
C. familiaris   MHDAIGFKSTLTGKNYTMEWYELFQLGNCTFPHLRPEMNAPFWCNQGAACFFEGIDDIHW 167
C. jacchus      MHDAIGFRSTLTGKNYTVEWYELFQLGNCTFPHLRPEMDAPFWCNQGAACFFEGIDDAHW 296
M. musculus     MHDAVGFRSTLTGKNYTIEWYELFQLGNCTFPHLRPDKSAPFWCNQGAACFFEGIDDKHW 158
R. norvegicus   MHDAIGFRSTLTGKNYTVEWYELFQLGNCTFPHLRPEVNAPFWCNQGAACFFEGIYDKHW 158
                ****:**:**** ****:******************: .**************** * **

H. sapien       KENGTLVQVATISGNMFNQMAKWVKQDNETGIYYETWNVKASPEKGAETWFDSYDCSKFV 284
P. troglodytes  KENGTLVQVATISGNMFNQMAKWVKQDNETGIYYETWNVKASPEKGAETWFDSYDCSKFV 284
M. mulatta      KENGTLVKVATISGNMFNQMAKWVKQDNETGIYYETWNVKASPEKGAETWFDSYDCSQFV 358
B. taurus       KENGTLVLVATISGGMFNRMAKWVKQDNETGIYYETWTVQASPERGAERWFESYDCSKFV 234
O. aries        KENGTLVLVATISGGMFNKMAKWVKQDNETGIYYETWTVQASPKKEAEKWFESYDCSKFV 237
C. familiaris   KENGTLVLVATISGNTFNQMAKWVKRDNETGIYYETWTVQASPTKGAETWFESYDCSKFV 227
C. jacchus      KENGTLVQVATISGNMFNKMAKWMKQDNETGIYYETWTVKASPEKGAETWFDSYDCSKFV 356
M. musculus     KENGTLSVVATISGNTFNKVAEWVKQDNETGIYYETWTVRAGPGQGAQTWFESYDCSNFV 218
R. norvegicus   KENGTLSLVATVSGNTFNKVAEWVKQDNETGIYYETWTVRASPGKGAQTWFESYDCSNFV 218
                ******  ***:**. **::*:*:*:*:***********.*:*.* : *: **:*****:**
```

```
H. sapien         LRTFNKLAEFGAEFKNIETNYTRIFLYSGEPTYLGNETSVFGPTGNKTLGLAIKRFYYPF 344
P. troglodytes    LRTFNKLAEFGAEFKNIETNYTRIFLYSGEPTYLGNETSVFGPTGNKTLGLAIKRFYYPF 344
M. mulatta        LRTFSKLAEFGAEFKNIETNYTRIFLYSGEPTYLGNETSVFGPTGNKTLGLAIKRFYYPF 418
B. taurus         LRTYEKLAELGADFKKIETNYTRIFLYSGEPTYLGNETSVFGPTGNKTLALAIKKFYYPF 294
O. aries          LRTYEKLAELGADFKKIETNYTRIFLYSGEPTYLGNETSVFGPTGNKTLALAIKKFYYPF 297
C. familiaris     LRTYKKLAELGAEFKKIETNYTRIFLYSGEPTYLGNETSIFGPTGNKTLALAIKRFYYPF 287
C. jacchus        LRTYNKLAEFGAEFKKIETNYTRIFLYSGEPTYLGNETSVFGPTGNKTLALAIKKFYYPF 416
M. musculus       LRTYKKLAEFGTEFKKIETNYTKIFLYSGEPIYLGNETSIFGPKGNKTLALAIKKFYGPF 278
R. norvegicus     LRTYEKLAEFGTDFKKIETNYTKIFLYSGEPIYLGNETSIFGPKGNKTLALAIKKFYGPF 278
                  ***:.****:*::**:******:******** *******:***.*****.****:** **


H. sapien         KPHLPTKEFLLSLLQIFDAVIVHKQFYLFYNFEYWFLPMKFPFIKITYEEIPLPIR-NKT 403
P. troglodytes    KPHLPTKEFLLSLLQIFDAVIVHRQFYLFYNFEYWFLPMKFPFIKITYEEIPLPNR-NKT 403
M. mulatta        KPHLSTKEFLLNLLQIFDAVIVHREFYLFYNFEYWFLPMKFPFIKITYEEIPLPNK-NKT 477
B. taurus         KPHLSTKEFLLSLLQIFDAVVIHREFYLFYNFEYWFLPMKYPFIKITYEEIPLPNRKNRT 354
O. aries          KPHLSTKEFLLSLLQIFDAVVIHREFYLFYNFEYWFLPMKSPFIKITYEEIPLPNRKNRT 357
C. familiaris     KPHLSTKEFLLSILQIFDAVIIHREFYLFYNFEYWFLPMKFPFIKITYEEIPLPKR-NET 346
C. jacchus        KPHLSTKEFLLSLLQIFDAVIVHRQFYLFYNFEYWFLPMKFPFIRITYEEIPLPKR-NKT 475
M. musculus       RPYLSTKDFLMNFLKIFDTVIIHRQFYLFYNFEYWFLPMKPPFVKITYEETPLPTR-HTT 337
R. norvegicus     KRYSSTKDFLLNFLKIFDTVIMHREFYLFYNFEYWFLPMKPPFVKITYEETPLPTQ-HTT 337
                  : : .**:**:.:*:***:*::*::*************** **::***** *** : : *

H. sapien         LSGL-------- 407
P. troglodytes    LSGL-------- 407
M. mulatta        LSGL-------- 481
B. taurus         LSGL-------- 358
O. aries          LSGL-------- 361
C. familiaris     LSGL-------- 350
C. jacchus        LSARHGSSHL-- 485
M. musculus       FTDL-------- 341
R. norvegicus     FTDL-------- 341
```

Supplementary Figure S2- Sequence alignment of CLN5. Alignment of mammalian CLN5 protein sequences. Yellow highlights indicate conserved N-glycosylation sites among different species while the green highlight indicates the N-glycosylation site corresponding to human N401, which is not conserved in rodents such as *M. musculus* and *R. norvegicus*. Sequences used in this alignment: *H. sapien* NP_006484, *P. troglodytes* XP_509687, *M. mulatta* XP_001086029, *B. taurus* DAA23821, *O. aries* NP_001076064, *C. familiaris* NP_001011556, *C. jacchus* XP_002742639, *M. musculus* AAI41315, and R*. norvegicus* NP_001178618.

References

1. **Jalanko A, Braulke T** 2009 Neuronal ceroid lipofuscinoses. Biochim Biophys Acta 1793:697-709

2. **Santavuori P** 1988 Neuronal ceroid-lipofuscinoses in childhood. Brain Dev 10:80-83

3. **Xin W, Mullen TE, Kiely R, Min J, Feng X, Cao Y, O'Malley L, Shen Y, Chu-Shore C, Mole SE, Goebel HH, Sims K** 2010 CLN5 mutations are frequent in juvenile and late-onset non-Finnish patients with NCL. Neurology 74:565-571

4. **Kohlschutter A, Schulz A** 2009 Towards understanding the neuronal ceroid lipofuscinoses. Brain Dev 31:499-502

5. **Haltia M** 2006 The neuronal ceroid-lipofuscinoses: from past to present. Biochim Biophys Acta 1762:850-856

6. **Haltia M** 2003 The neuronal ceroid-lipofuscinoses. J Neuropathol Exp Neurol 62:1-13

7. **Kyttala A, Lahtinen U, Braulke T, Hofmann SL** 2006 Functional biology of the neuronal ceroid lipofuscinoses (NCL) proteins. Biochim Biophys Acta 1762:920-933

8. **Tyynela J, Palmer DN, Baumann M, Haltia M** 1993 Storage of saposins A and D in infantile neuronal ceroid-lipofuscinosis. FEBS Lett 330:8-12

9. **Jabs S, Quitsch A, Kakela R, Koch B, Tyynela J, Brade H, Glatzel M, Walkley S, Saftig P, Vanier MT, Braulke T** 2008 Accumulation of bis(monoacylglycero)phosphate and gangliosides in mouse models of neuronal ceroid lipofuscinosis. J Neurochem 106:1415-1425

10. **Kousi M, Lehesjoki AE, Mole SE** 2012 Update of the mutation spectrum and clinical correlations of over 360 mutations in eight genes that underlie the neuronal ceroid lipofuscinoses. Hum Mutat 33:42-63

11. **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 19:286-288

12. **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 13:2410-2420

13. **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 16:29-40

14. **Isosomppi J, Vesa J, Jalanko A, Peltonen L** 2002 Lysosomal localization of the neuronal ceroid lipofuscinosis CLN5 protein. Hum Mol Genet 11:885-891

15. **Mamo A, Jules F, Dumaresq-Doiron K, Costantino S, Lefrancois S** 2012 The Role of Ceroid Lipofuscinosis Neuronal Protein 5 (CLN5) in Endosomal Sorting. Mol Cell Biol 32:1855-1866

16. **Schmiedt ML, Bessa C, Heine C, Ribeiro MG, Jalanko A, Kyttala A** 2010 The neuronal ceroid lipofuscinosis protein CLN5: new insights into cellular maturation, transport, and consequences of mutations. Hum Mutat 31:356-365

17. **Lyly A, von Schantz C, Heine C, Schmiedt ML, Sipila T, Jalanko A, Kyttala A** 2009 Novel interactions of CLN5 support molecular networking between Neuronal Ceroid Lipofuscinosis proteins. BMC Cell Biol 10:83

18. **Wujek P, Kida E, Walus M, Wisniewski KE, Golabek AA** 2004 N-glycosylation is crucial for folding, trafficking, and stability of human tripeptidyl-peptidase I. J Biol Chem 279:12827-12839

19. **Aebi M, Bernasconi R, Clerc S, Molinari M** 2010 N-glycan structures: recognition and processing in the ER. Trends Biochem Sci 35:74-82

20. **Helenius A, Aebi M** 2004 Roles of N-linked glycans in the endoplasmic reticulum. Annu Rev Biochem 73:1019-1049

21. **Sleat DE, Wang Y, Sohar I, Lackland H, Li Y, Li H, Zheng H, Lobel P** 2006 Identification and validation of mannose 6-phosphate glycoproteins in human plasma reveal a wide range of lysosomal and non-lysosomal proteins. Mol Cell Proteomics 5:1942-1956

22. **Golabek AA, Kida E, Walus M, Wujek P, Mehta P, Wisniewski KE** 2003 Biosynthesis, glycosylation, and enzymatic processing in vivo of human tripeptidyl-peptidase I. J Biol Chem 278:7135-7145

23. **Sun Q, Westphal W, Wong KN, Tan I, Zhong Q** 2010 Rubicon controls endosome maturation as a Rab7 effector. Proc Natl Acad Sci U S A 107:19338-19343

24. **Wegner CS, Malerod L, Pedersen NM, Progida C, Bakke O, Stenmark H, Brech A** 2010 Ultrastructural characterization of giant endosomes induced by GTPase-deficient Rab5. Histochem Cell Biol 133:41-55

CHAPTER 3

DISCUSSION AND FUTURE PROSPECTIVES

DISCUSSION AND FUTURE PROSPECTIVES

The spectrum of mutations that have been recently identified to result in NCL has grown steadily and what was once considered to be a disease limited to a specific population is now known to have a diverse ethnic spread. Many genes have been discovered that are known to cause NCL. There is an immense effort going on to find ways to cure or prevent NCL. The progress of drug discovery for NCL has been hindered by the fact that function of most NCL proteins is unknown. It is imperative to characterize a protein before a deeper understanding of its function can be comprehended. CLN5 we known is one such protein whose function remains unknown. We have studied the importance of N-glycosylation on the transport and stability of CLN5 here. N-glycosylation is an important modification that takes place co/post-translationally. Although it is not the only kind of modification, it is still an important one as it is known to aid in the proper folding and functioning of proteins. In addition N-glycosylation also protects the protein from other hydrolases and unfavorable environments. We have shown that all the eight putative N-glycosylation sites are used *in vivo*. Since all the sites are used, it suggests that all sites play a role in either folding, stability and/or trafficking of CLN5, although it is possible that one or more N-glycosylation sites are paramount for CLN5 function and/or transport. This study will help in understanding the role of CLN5 in NCL better. It will aid later studies on CLN5 as it expands on some of the basic characterization of CLN5.

Mass spectrometry and bioinformatics studies have predicted that, of the eight potential N-glycosylation sites, the last three - Asn320, 330 and 401 potentially have a M6P tag. M6P is essential for the recognition by M6PR to transport cargo to the endosomes. Some LSD

result due to faulty transport of proteins. I-cell disease is one of them where there a defect in the enzyme that is involved in the addition of M6P tag. In I-cell disease, lysosomal hydrolases and other lysosomal proteins that are normally transported via the M6P pathway to the lysosomes end up being secreted out of the cell as they do not have an M6P tag anymore. Of the three predicted N-glycosylation sites with an M6P tag in CLN5, our studies indicate that Asn401 to be a prime candidate in the transport of CLN5 by M6P pathway. We feel Asn401 to be the prime candidate because mutating Asn320 to Gln results in the protein being retained in the ER and as such it is difficult to assess its role in M6P pathway. N330Q mutant was able to locate to the lysosomes, but it was also seen to be retained in the ER. Thus, Asn330 does not seem to play an integral role in the M6P mediated transport of CLN5 to the lysosomes. Asn401 on the other hand, when mutated to Gln caused it to be retained mostly in the Golgi and was not observed in the ER or lysosomes. CHX chase studies also showed that the N401Q mutant gets secreted out into the media very early and is more prominent than CLN5 wt. The loss of M6P tag is known to result in lysosomal proteins being secreted out in I-cell disease, indicating that Asn401 might play a similar role in targeting CLN5 to the endosomes/lysosomes. Although studies have shown CLN5 to utilize M6P independent pathways to reach the lysosomes, these studies were done in mice and interestingly, mouse CLN5 lacks the Asn401 N-glycosylation site. It is possible that there are other mechanisms that compensate for the loss of Asn401 site in mouse. From our studies the importance of Asn401 in human CLN5 transport cannot be ignored. Future studies can focus on elucidating the role of Asn401 in the transport of CLN5 to the endosomes via the M6P pathway. It is not known if M6P pathway is the only pathway utilized by human cells for the trafficking of CLN5. One way to elucidate the role of M6P in CLN5

trafficking would be to use M6PR knockout cells or use siRNA against M6PR and look at the subcellular localization of human CLN5. It is possible that in the absence of M6PR, CLN5 could utilize other adaptors/mechanisms/pathways to reach the lysosomes (e.g., LIMP and Sortilin). A similar modus operandi of knocking out or silencing these transport adaptors along with M6PR can be employed to look at CLN5 localization.

Among the known CLN5 mutations that cause NCL, N192S is on one of the N-glycosylation sites. This mutant localizes to the lysosome. Our studies also show that abrogating the N-glycosylation site at Asn192 does not affect the localization of CLN5 and it ends up in the lysosome. The fact that ablation of this site is known to cause NCL but does not affect the localization is an important indication that this site is essential for the proper functioning of CLN5. It does not seem to have a very important role in the transport or proper folding of CLN5. This mutation (N192S) can be used as a tool to study the functioning of CLN5 in the future, along with the other mutations of N-glycosylation in CLN5 that are able to reach the lysosomes (N192Q and N227Q). Since CLN5 is known to interact with other NCL proteins, it would be interesting to look at the interaction (or loss of) of CLN5 N192S mutant along with N192Q and N227Q. It is possible that abrogation of N-glycosylation at Asn192 renders CLN5 unable to bind to its normal partners resulting in NCL. Co-IP experiments of CLN5 wt and N192S mutant will help in clarifying this point.

CLN5 is a highly glycosylated protein and removal of the glycosylation (Endo H or PNGase F digestion) causes the size of the CLN5 protein to drop drastically. We have seen that PNGase F digested CLN5 runs slightly higher than the tunicamycin treated CLN5. This hints at

the presence of other modifications on CLN5. CLN5 could be O-glycosylated, phosphorylated, sulfated or lipoylated. Digesting CLN5 with O-glycosidase alone or in combination with PNGase F would be a means of figuring the extra modification on CLN5. Another outlet would be to treat CLN5 with a general phosphatase alone or in combination with PNGase F for the same purpose as above.

Clearly N-glycosylation plays an important role in the proper functioning, transport and stability of CLN5 and this study corroborates these facts. This study adds to the basic understanding of CLN5. Having said that, there is still immense scope for understanding CLN5 protein and how it functions in NCL disease.