Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0148.v1

Peptide: N-glycanase 1 and its relationship with congenital disorder of deglycosylation

Xiangguang Miao[1], Jin Wu[2,3], Hongping Chen[4,#], Guanting Lu[2,#]

1. Queen Mary School, Nanchang University, No. 1299 Xuefu Avenue, Honggutan New District, Nanchang, 330036, China

2. Laboratory of Translational Medicine Research, Department of Pathology, Deyang Key Laboratory of Tumor Molecular Research, Deyang People's Hospital, No. 173 First Section of Taishanbei Road, Jingyang District, Deyang, 618000, China

3. Department of Molecular & Cellular Biology, Roswell Park Comprehensive Cancer Center, Buffalo, NY, USA.

4. Jiangxi Key Laboratory of Experimental Animals, Nanchang University, Nanchang, 330006, China.

Correspondence:

Guanting Lu, Laboratory of Translational Medicine Research, Department of Pathology, Deyang Key Laboratory of Tumor Molecular Research, Deyang People's Hospital, No. 173 First Section of TaishanBei Road, Jingyang District, Deyang, 618000, China. guantlv@126.com; +86-18801474087.

Hongping Chen, Jiangxi Key Laboratory of Experimental Animals, Nanchang University, Nanchang, 330006, China. jxchp2000@126.com.

© 2022 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0148.v1

Abstract

The cytosolic PNGase (peptide:N-glycanase; Png1 in yeast; NGLY1/Ngly1 in human/mice), also known as peptide-N4-(N-acetyl-beta-glucosaminyl)-asparagine amidase, is a well-conserved deglycosylation enzyme (EC 3.5.1.52) which catalyzes the non-lysosomal hydrolysis of an N(4)-(acetyl-β-D-glucosaminyl) asparagine residue into N-acetyl-β-D-glucosaminylamine and a peptide containing an aspartate residue. This enzyme (NGLY1) plays essential roles in clearance of misfolded or unassembled glycoproteins through a process named ER-associated degradation (ERAD). Accumulating evidence also points out that NGLY1 deficiency can cause an autosomal recessive human genetic disorder associated with abnormal development and congenital disorder of deglycosylation. In addition, the loss of NGLY1 can affect multiple cellular pathways, including but not limited to NFE2L1 pathway, Creb1/Atf1-AQP pathway, BMP pathway, AMPK pathway, and SLC12A2 ion transporter, which might be the underlying reasons for a constellation of clinical phenotypes of NGLY1 deficiency. The current comprehensive review indeed uncovers the detailed NGLY1's structure and its important roles for participation in ERAD, involvement in CDDG and potential treatment for NGLY1 deficiency.

Keywords

N-glycosylation; NGLY1; ER associated degradation process; congenital disorder of deglycosylation; NFE2L1

Introduction

In living eukaryotes, secretory proteins are translocated into the endoplasmic reticulum (ER) lumen via signal recognition particles (SRPs) and subjected to a quality control process. During this process, proteins are modified and sorted in ER based on their three-dimensional conformation. The correctly folded proteins are then transferred to the Golgi apparatus for maturation and rerouted to other cellular compartments via transport vesicles. The misfolded or unassembled proteins are subjected to degradation through a process known as ER-associated degradation (ERAD) [1].

The cytosolic PNGase (peptide: N-glycanase; Png1 in yeast; NGLY1/Ngly1 in hu-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0148.v1

man/mice), also named as peptide-N4-(N-acetyl-beta-glucosaminyl)-asparagine amidase, is a well-conserved deglycosylation enzyme (EC 3.5.1.52) which catalyzes the non-lysosomal hydrolysis of an N(4)-(acetyl-β-D-glucosaminyl) asparagine residue into N-acetyl-β-D-glucosaminylamine and a peptide containing an aspartate residue [2]. NGLY1 plays essential roles in clearance of misfolded glycoproteins via releasing intact N-glycans from N-glycosylated proteins. Accumulating evidences point out that NGLY1 deficiency can cause an autosomal recessive human genetic disorder with abnormal development, indicating its critical functions for the normal development of mammals [3; 4]. Current review mainly focuses on NGLY1's structure, participation in ERAD, involvement in CDDG and potential treatment for NGLY1 deficiency.

Briefings about the discovery of peptide: N-glycanase

The glycopeptidase partially purified from the seed emulsion of a fruit plant (almond) was first reported in 1977 to show its enzymatic activity to cleave β-aspartylglycosylamine linkages in glycopeptides [5]. Later, this glycopeptidase was discovered in several other plants, including jack bean (*Canavalia ensiformis*) [6], split pea (*Pisum sativum*) [7], Arabidopsis thaliana [8], lentil (Lens culinaris), pinto bean, lima bean, barley and wheat [7]. This implied that the existence of this specific glycopeptidase might be ubiquitous in plant kingdom. In 1978, some basic enzymatic characters of this new glycopeptidase such as Km value, optimum pH, and inhibitors were revealed [9]. Since the enzyme could only hydrolyzed N-glycopeptides longer than 3 amino acids [9], it was also called as peptide: N-glycosidase (PNGase). It's reported in 1981 that PNGase could highly efficiently cleave short ovalbumin and bromelain N-glycopeptides modified with high content of mannose and complex oligosaccharides [10]. Interestingly, it's discovered at the same year that the carbohydrate removal of pepsin-digested glycoproteins by PNGase was almost three times higher than the intact untreated proteins [11]. This implied that the complex oligosaccharides might hinder the accessibility of PNGase to the cleavage site of the target glycoproteins under their native state. Treating glycoproteins with heat in sodium dodecyl sulfate (SDS) or high concentrations of chaotropic salts such as NaSCN and $NaClO_4$ with disulfide bond reducing agent β-mercaptoethanol (BME) to denature the native structures, more N-glycans were released from the denatured and unfolded proteins upon treatment with PNGase [12]. In

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0148.v1

1984, the enzymatic activity of PNGase was discovered in the Gram-negative bacterial pathogen for meningitis and septicaemia, Flavobacterium meningosepticum [13]. This was the first time for the discovery of PNGase in prokaryotes. More importantly, in 1989, free di-N-acetylchitobiose based oligosaccharides were detected in the extract of eggs of trout (*Plecoglossus altivelis*) [14], dace (*Tribolodon hakonensis*) and flounder (*Paralichthys olivaceus*) [15]. Since di-N-acetylchitobiose was mainly produced from N-linked glycoproteins by PNGase, it's quite reasonable to infer the presence of PNGase in the animal kingdom. This led to the discovery of the Peptide: N-Glycosidase in the embryos of Madaka fish (*Oryzias latipes*) in 1991 [16]. Two years later, PNGase was discovered in mouse (L-929, BALB-3T3 and P3X63-Ag8.U1) and human (TIG-3S) cell lines [17]. In 1997 and 1998, PNGase was identified in two fungi, *Aspergillus tubigensis* [18] and *Saccharomyces cerevisiae* [19], respectively. Up to now, the existence of PNGase has been found in the plant, prokaryotes, animal and fungi kingdoms (Figure 1). This PNGase was named as Png1 in Arabidopsis thaliana, Png1 in yeast, png-1 Caenorhabditis elegans, Pngl in fruit fly, Ngly1 in mice and rat, and NGLY1 in human. For readability, the gene is termed as NGLY1/Ngly1 in the whole article.

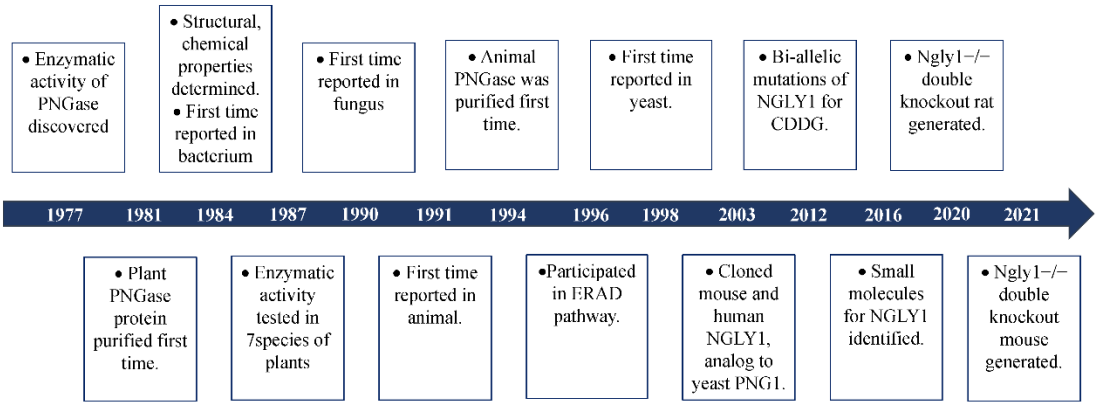

Figure 1. Brief history of the discovery of PNGase.

Based on experimental and bioinformatic analysis, Ngly1 was universally expressed in almost all the living cells (Figure 2A). It has been verified that Ngly1 played an essential role in the degradation of misfolded glycoproteins. Loss-of-functional (LoF) mutations of Ngly1 might be detrimental to the deglycosylation process. In 2012, bi-allelic compound mutations of NGLY1 were identified by whole-exome sequencing (WES) in a boy with congenital anomalies and/or intellectual disabilities [3]. This was the first time to associate NGLY1 with a specific disorder. Since the disorder was

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0148.v1

caused by pathologic mutations in the specific enzyme responsible for deglycosylation, therefore, the disease was named as congenital disorder of deglycosylation (CDDG) or NGLY1-congenital disorder of deglycosylation (NGLY1-CDDG; OMIM#615273). With the discovery of mutations in NGLY1 for CDDG [3; 4; 20; 21; 22], more and more researches are concentrated on the molecular mechanisms of NGLY1 in the process of deglycosylation (Figure 1).

General structure of NGLY1

In human, NGLY1 gene (NM_018297) contains 12 exons spanning about 70 kb on chromosome 3p24.2. The encoded protein is composed of 654 amino acids with three functional domains, namely, a N-terminal PUB domain, a central transglutaminase-like domain and a C-terminal PAW domain predicted by AlphaFold (Figure 2A) [23]. A short intrinsically disordered region was identified ranging from 113 to 164 between the PUB and transglutaminase-like domain by IUPred2A [24]. Disorder Enhanced Phosphorylation Predictor (DEEP) revealed a phosphorylation hotspot in this intrinsically-disordered region (Figure 2B), indicating its importance for the function of NGLY1.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0148.v1

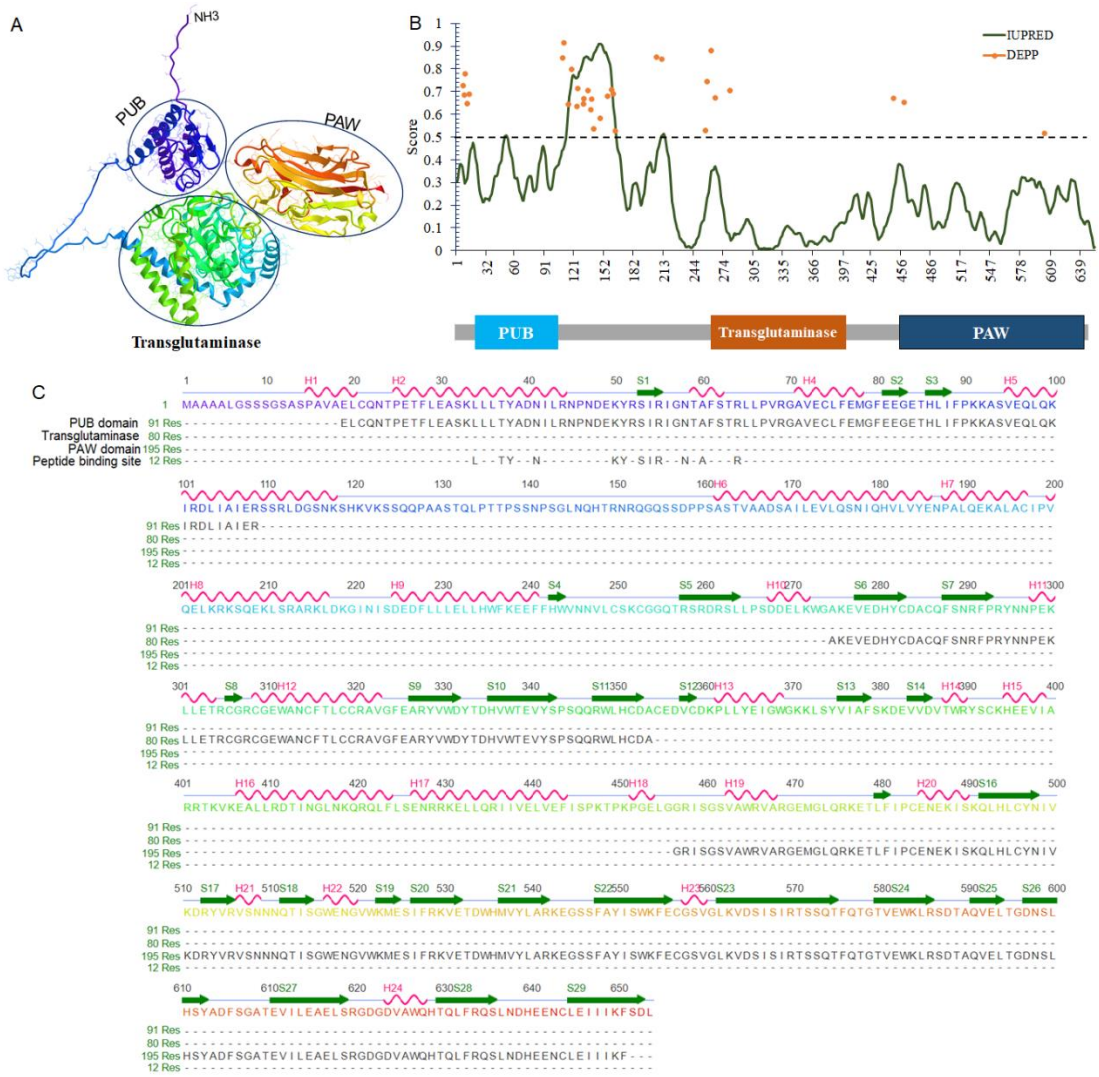

Figure 2. Structural analysis of NGLY1. A) Three-dimensional structure predicted by AlphaFold; B) Intrinsically-disordered region of NGLY1; C) Secondary structure of NGLY1.

N-terminal PUB domain

The 91 aa-PUB (*P*NGase/*UB*A or *UB*X-containing protein) domain, located at the N-terminal (aa19-109) of NGLY1, was first identified through bioinformatics analysis [25; 26]. According to the AlphaFold Protein Structure Database, this PUB domain consists of a bundle of five α-helices that pack onto a short three stranded anti-parallel β-sheets (Figure 2C) [27]. Based on evolutionary conservation analysis, PUB domain was existed in almost all the members in animals, except for snakes and nematodes. Considering the absence in certain species, it's quite reasonable to infer that this domain

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0148.v1

may not be directly involved in the catalytic activity of NGLY1, but rather played other important roles. The ubiquitin-associated (UBA) domain was reported in many ubiquitin-regulatory proteins (UBX) being involved in the ubiquitination pathway or containing regions homology to ubiquitin itself, such as CBL, CBLB, MARK3, RAD23A, RAD23B, and UBA2 and UBXN1 [28]. Since the most conserved residues in the UBA domain were non-polar amino acids to stabilize the secondary α-helical structure, it's indicated that the UBA domain was unlikely to be directly involved in phosphorylation or ubiquitination, but in mediating protein-protein interactions [25; 29; 30; 31]. The PUB motif could bind to the adaptor protein p97 (also called VCP, Cdc48) which was located on the endoplasmic reticulum (ER) membrane [29; 32], which provided NGLY1 with a molecular platform for the de-glycosylating process toward the retrotranslocated misfolded or unassembled proteins [33]. It could also bind with other proteins functioning in the ER-associated degradation (ERAD) pathway, such as DERL1 [34].

Central transglutaminase-like core domain

Since containing the transglutaminase-like core domain, NGLY1 was categorized as a member in the transglutaminase superfamily featuring the Cys, His, and Asp catalytic triad, which was essential for the enzymatic function [2]. The core domain could also be subdivided into transglutaminase-like region, zinc-binding motif, and RAD23/HR23 binding motif [35; 36]. The transglutaminase sequence was responsible for breaking the β-aspartylglycosylamine linkage between the innermost N-acetylglucosamine (GlcNAc) and the Asn residue (N-X-S/T) of the target N-glycoprotein. After surveying the NGLY1 protein sequence, there are four CXXC motifs around or in the transglutaminase-like region. Since the CXXC sequence could bind with Zn^{2+} to form zine-finger motifs, Zn^{2+} might be important for the enzymatic activity. It had been reported that several divalent metal ions, such as Zn^{2+}, Mg^{2+}, Co^{2+}, and Cu^{2+} could greatly increase the N-glycanase velocity [37]. However, in most filamentous fungal species such as Neurospora crassa, Cys and His of the catalytic triad in NGLY1 were mutated leading to complete loss of enzymatic activity [38]. This indicated that the N-glycanase enzymatic activity was not essential for some organisms. The RAD23/HR23 binding motif rendered NGLY1 with the ability to bind with Rad23 (RAD23A and RAD23B in human), thus taking part in the ubiquitin-proteasome pathway [39; 40]. Rad23 could

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0148.v1

also tightly bind with xeroderma pigmentosum group C (XPC) and participated in the nucleotide excision repair (NER) pathway [41; 42]. Therefore, the central transglutaminase core domain of NGLY1 could not only hydrolyze the β-aspartylglycosylamine linkage, but also act as adaptor to bind with components in the proteasome or NER pathway.

C-terminal PAW domain

Compared with the protein structures in different species, a PAW domain was located at the C-terminus of NGLY1. It was named for 'domain present in *P*NGases *a*nd other *w*orm proteins'. The PAW domain has a β-sandwich architecture composed of 2 layers containing antiparallel β strands, and short helices (Figure 2C). The PAW domain was reported to bind with the high content mannose moieties of N-linked oligosaccharide chains [43]. The interaction between PAW domain and the mannose could increase the affinity of NGLY1 for N-glycans to promote the enzymatic activity. Combination of transglutaminase core domain and PAW domain contributes to the oligosaccharide-binding specificity of NGLY1 for the N-linked glycans with high mannoses.

Evolutionary relationships of NGLY1

The protein sequences of NGLY1 from different species were recruited from the NCBI's GenBank and aligned by CLUSTALW. The evolutionary phylogeny was constructed using the Minimum Evolution method (Figure 3A) [44]. The structures of these NGLY1 from different species were strongly conserved during evolution (Figure 3B). This implied the importance of the protein N-glycosylation and de-N-glycosylation system for life, from simple single cells to the most sophisticated mammals. It's clearly shown in animal kingdom that except for tiger snake (Notechis scutatus) and nematode (Caenorhabditis elegans), other NGLY1 contains three different domains, a PUB domain, a transglutaminase (TG) domain and a PAW domain. The Ngly1 protein of tiger snake lacks the PUB domain, but has a BAR (Bin/Amphiphysin/Rvs) motif in the middle of TG and PAW domains. It has been reported that the BAR domain possesses the ability to bind with cellular membranes [45; 46; 47; 48], this might substitute for the function of PUB domain to provide a platform for the de-glycosylating process. The *C. elegans* Ngly1 (png-1) contains a TRX (thioredoxin) domain, instead of the PUB motif. The TRX domain confers Ngly1 with an extra thioredoxin ability [49] and makes Ngly1 a unique bi-functional protein possessing two enzyme activities [50]. It's worth noting

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0148.v1

that instead of lacking the specific PUB domain, the Ngly1 in chimpanzee (Pan troglodytes) and Pacific oyster (Crassostrea gigas) were predicted to have one more FRG2 (cl21160) and Nucleo_P87 (pfam07267) domain, respectively. However, the functions of these two domains were unknown. Experiments should be performed to reveal if these types of Ngly1 were tri-functional proteins.

In the fungi and plants kingdoms, the Ngly1 only contained the transglutaminase (TG) domain, without the PUB and PAW domains [51]. However, the length of plant Ngly1 was almost two times longer than those of fungi. It implied that the plant Ngly1 might possess more complex functions than fungi. In the dicot plant, Arabidopsis thaliana, the Ngly1 has a β-grasp ubiquitin-like fold (cl28922) at the N-terminal. In the monocot plant, rice (Oryza sativa Japonica) Ngly1 has a β-grasp ubiquitin-like fold (cl28922) at the N-terminal and a F5/8 type C domain (pfam00754) at the C-terminal.

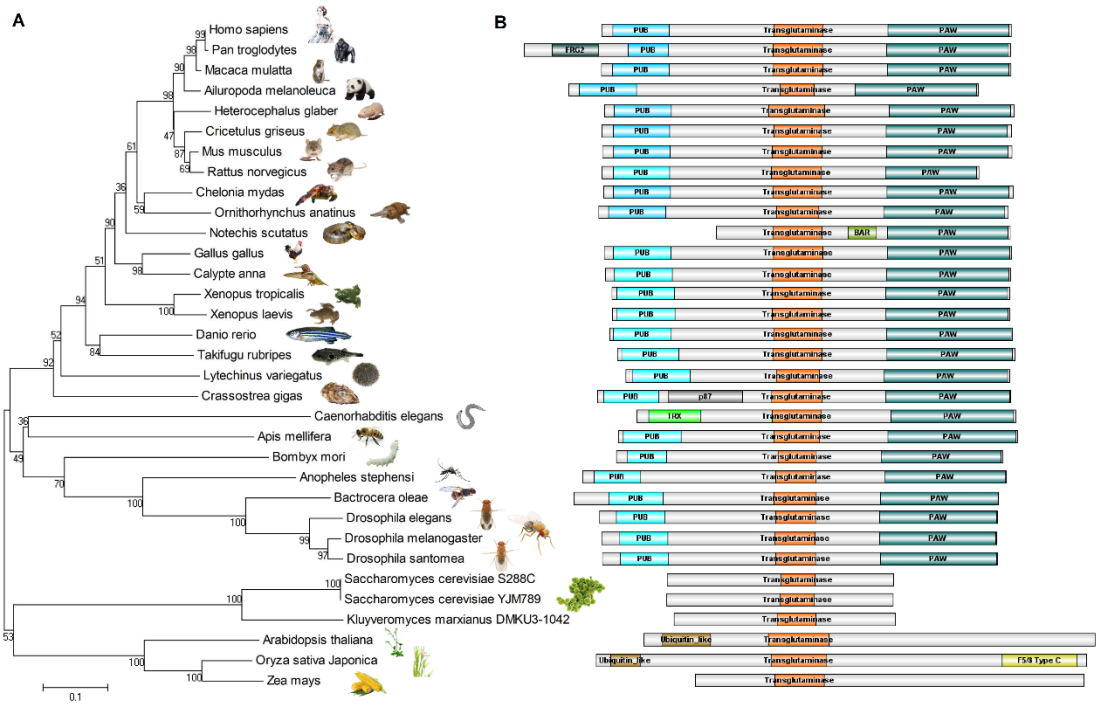

Figure 3. Evolutionary relationships of NGLY1 in 33 taxa. A) Evolutionary phylogeny; B) Protein structures of Ngly1.

Roles of NGLY1 in ERAD pathway

It had been reported that only the N-linked glycans on short peptides, pepsin-digested, denatured or misfolded glycoproteins could be highly efficiently removed by Ngly1 [11; 12]. This implied that the complex N-linked oligosaccharides provide a steric hindrance effect to inhibit the accessibility of Ngly1 to the hydrolyzation site of the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0148.v1

intact target proteins. In living cells, most proteins were glycosylated in the ER lumen by a series of glycosylases to prevent the abnormal aggregation and promote the transportation to their target cellular compartments [52; 53]. Even under normal physiological conditions, some rapidly synthesized or mutated proteins could be improperly folded (or called misfolded). The accumulation and aggregation of misfolded proteins would induce cellular toxicity to the cell if left unchecked [54; 55]. It's well known that the toxic amyloid-like aggregates of β-amyloid (APP), tau (MAPT), α-synuclein (SNCA or PARK1), and islet amyloid polypeptide (IAPP) were linked with neurodegeneration in Alzheimer's disease (AD), frontotemporal dementia (FTD), Parkinson's disease (PD) or Type 2 diabetes (T2D) [56]. Therefore, timely clearance of the misfolded or misassembled proteins is essential for the physiological cellular homeostasis and cell growth. Fortunately, a quality control system in the ER was evolved to pick misfolded oligopeptides out of the properly folded proteins based on their tertiary conformation [57]. The picked-out misfolded proteins were retrotranslocated into the cytosol and then polyubiquitinated, de-glycated and targeted to the 26S proteasome for degradation, this distinctive process was called ER-associated degradation pathway (ERAD) (Figure 4-upper panel) [1; 58].

Misfolded glycoproteins were transferred from ER lumen to cytosol by a protein complex consisted of SEL1L1 (sel-1), SYVN1 (hrd1) and DERL1 (derlin-1). In cytosol, the VCP-UFD1L1-NPLOC4 (p97-Ufd1-Npl4) AAA ATPase complex was interacted with SEL1L1 and SYVN1. With the energy-providing multiprotein ATPase complex (p97-Ufd1-Npl4), misfolded proteins dedicated for degradation were efficiently retrograded into the cytosol from ER lumen [59; 60]. In cytosol, a VCP-binding multiprotein complex Cullin-RING Ligase 1 (CRL1) was formed by a scaffolding cullin protein CUL1, RBX1/Roc1, SKP1 and substrate-binding adaptor F-box proteins (such as FBXO2/Fbs1 and FBXO6/Fbs2), therefore, the complex also named as SCF (SKP1-cullin-F-box) [61; 62]. Upon retrograded into cytosol, the misfolded target proteins with exposed Man3-9GlcNAc2 glycans were bound by the FBXO2 or FBXO6 and polyubiquitinated by the ubiquitin ligase in the SCF complex [62; 63; 64; 65]. Subsequently, NGLY1 was hold tightly onto VCP/p97 of the p97-Ufd1-Npl4 complex via its N-terminal PUB domain [29]. Its PAW domain grip onto the mannose moieties of the polyubiquitinated misfolded proteins and the transglutaminase-like (TG) domain

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0148.v1

cleaved the β-aspartylglycosylamine linkage between the innermost N-acetylglucosamine (GlcNAc) and the Asn residue (N-X-S/T), releasing a free Man3-9GlcNAc2 and a de-glycated protein with an Asp (D) instead of the Asn (N) residual. The processed proteins with Ds were transferred to the 26S proteasome which was linked with NGLY1 via the HR23 complex (RAD23A and RAD23B) for thorough degradation. The free GlcNAc oligosaccharides were recycled [66] or thoroughly hydrolyzed by a series of enzymes including ENGase (endo-β-N-acetylglucosaminidase) [55; 67].

However, when NGLY1 was deleted or mutated with bi-allelic loss-of-functional mutations leading to NGLY deficiency, the process of ER-associated degradation (ERAD) was interrupted after ubiquitinated by SCF ligase complex. The Man3-9GlcNAc2 glycans in the misfolded proteins could not be released by NGLY1, the glycosidic bond between the two GlcNAcs was cut by ENGase which released a free Man3-9GlcNAc and a protein with a N-GlcNAc. With accumulation of the N-GlcNAc misfolded proteins, they might be intertwined together to form insoluble aggregates in the cytosol. The aggregates could inhibit the protein degradation function of the 26S proteasome complex, thus rendered cellular toxic effect leading to apoptosis of the cells (Figure 4-Lower panel).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0148.v1

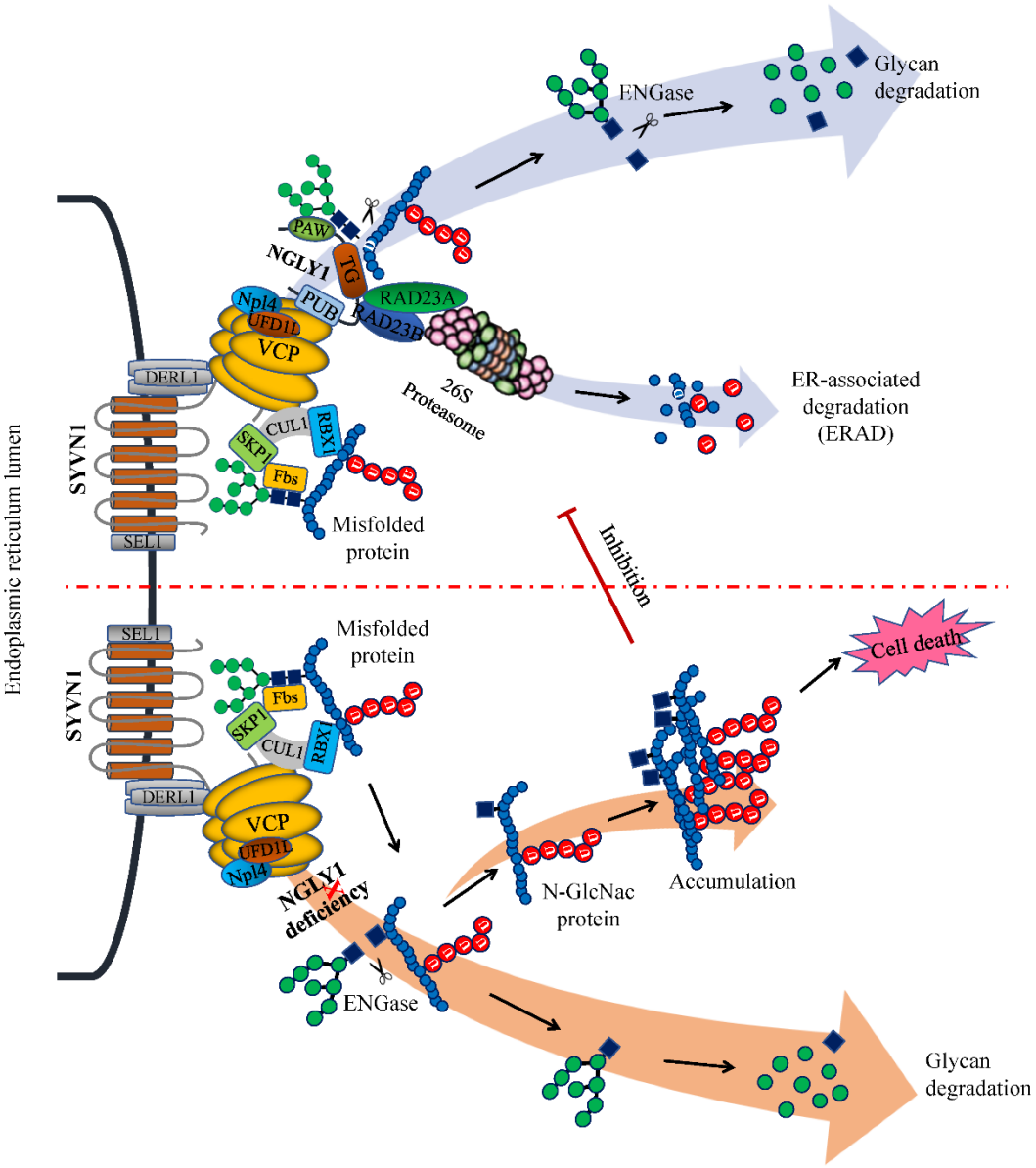

Figure 4. ER-associated degradation pathway (ERAD) involving NGLY1. Upper panel) ERAD under normal physiological state; Lower panel) ERAD with NGLY1 deficiency.

Different functional pathways participated by NGLY1

It was widely accepted that NGLY1 played an important role in the ER-associated degradation pathway to remove misfolded glycosylated proteins which were retrotrans-located from ER lumen to the cytosol [1; 4; 55]. Under condition of NGLY1 deficiency, the misfolded glycoproteins would be accumulated to form insoluble aggregates [4] which could inhibit the activity of 26S proteosome and lead to ER stress and cell death [68]. It seemed quite likely that NGLY1 might be involved in a general mechanism to deglycosylate misfolded glycoproteins through ERAD pathway. However, it had been

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0148.v1

reported that some proteins could be degraded regardless of the glycosylation status [69; 70]. In cells with NGLY1 deficiency from Drosophila melanogaster [71] and rat [72], mouse [73] and human [74], no evidences of ER stress were detected which was often observed in cells with impaired ERAD function. This indicated that NGLY1 might played diversified roles in living cells, in addition to involving in ERAD.

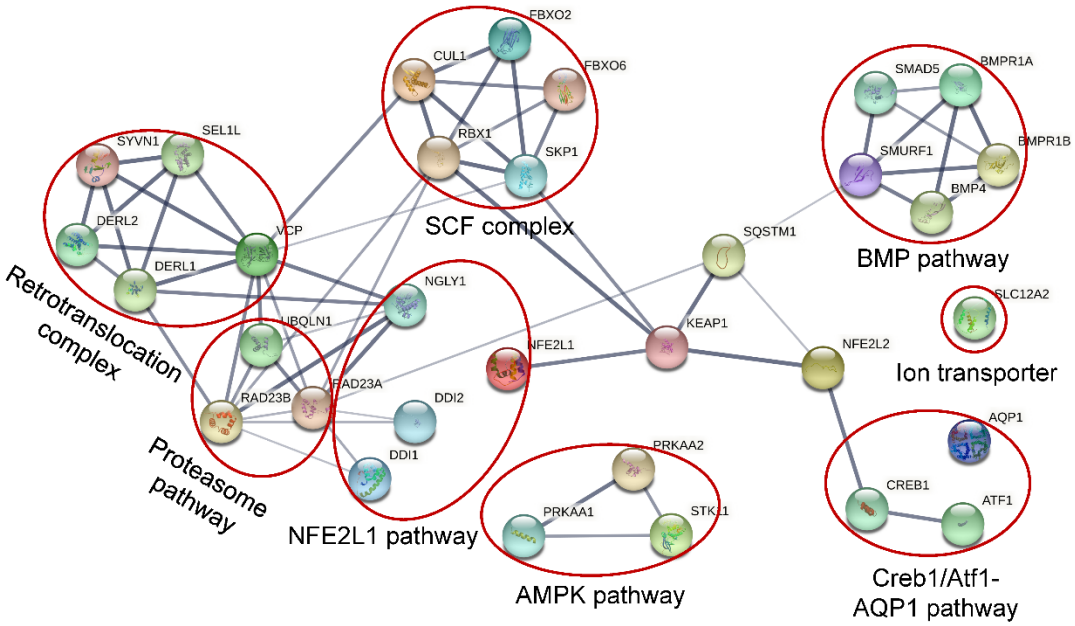

Figure 5. Protein network involving NLGY1 by STRING.

NFE2L1 pathway: NFE2 like bZIP transcription factor 1 (NFE2L1, also called NRF1) is a transcription factor belonging to the CNC-bZIP family [75]. NFE2L1 is involved in regulation of many cellular functions, such as oxidative stress response, differentiation, inflammatory response, and metabolism. NFE2L1 was ubiquitously expressed and could be induced by cellular stresses such as oxidative stress, ER stress, and inflammation. Under normal physiological state, the N-glycosylated non-active NFE2L1 was retrotranslocated into cytosol for ERAD-dependent degradation by the proteasome [76]. Under cellular stimuli for inhibition of the proteasome function, the retrotranslocated NFE2L1 was escaped the ERAD and the N-glycans were removed by non-ER bound NGLY1 to make the amino acid transition from Asn (N) to Asp (D). The deglycosylated NFE2L1 was further cleaved at the Leu104 residue by DNA damage inducible 1 homologs (DDI 1 or DDI 2) to release the ER-bound NFE2L1 from the ER membrane [77; 78]. The processed NFE2L1 (p110) was then entered into the nucleus to activate expression of a subset of proteasome subunits to promote the proteasome function to alleviate cellular stresses [75]. However, in NGLY1 deficient cells, the N-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0148.v1

linked glycans could not be released which inhibited the cleavage and activation of NFE2L1. It had been identified that NFE2L1 was highly expressed in the brain, heart, kidney, skeletal muscle, and fat [79]. Disruption of NFE2L1 pathway caused by NGLY1 mutations might be related with the neurological, renal, skeletal and ophthalmological phenotypes of patients with NGLY1 deficiency.

Creb1/Atf1-AQP pathway: It has been reported that regardless of the N-glycanase enzymatic activity, NGLY1 deficiency could decrease transcriptionally the levels of multiple aquaporins (AQPs) by 50%-60% in human and mouse cells indirectly through transcription factors Atf1/Creb1 [73]. Since aquaporins function as water transmembrane transporter, their decrease could explain partly the reason of poor or absent tear production, dry mouth, reduced saliva production [80] and constipation [81].

BMP pathway: In Drosophila, loss of Ngly1 could result in developmental midgut defects, which were similar with the deficiency of BMP signaling [82]. Later, Ngly1 was found to colocalize with endoplasmic reticulum via VCP (p97) to promote the retrotranslocation and de-glycosylate the misfolded BMP4 for proteasome degradation, which could increase the efficient traffic of properly-folded BMP4 to its target compartment through secretary pathway [83]. It had been reported that BMP4 was essential for mesoderm development, limb formation, tooth development, bone induction, nephric duct formation, renal system segmentation and aortic valve morphogenesis [84; 85; 86]. Mutations of BMP4 could result in eye and brain developmental anomalies [84; 87], which were overlapped with NGLY1 deficiency.

AMPK pathway: In some NGLY1 deficiency patients, altered muscle and liver mitochondrial amount, function and impaired physiology were identified, which could be rescued by restoration of NGLY1 expression, which confirmed the direct relationship of NGLY1 to mitochondrial function [88]. Recently, it has been reported that NGLY1 deficiency could severely reduce the expression and phosphorylation of AMP-activated protein kinase α (AMPKα) in Drosophila larval intestine, mouse embryonic fibroblasts and patient-derived fibroblasts, leading to energy metabolism defects, impaired gut peristalsis, failure to empty the gut, and animal lethality [89]. The reduced AMPKα expression observed in NGLY1 deficiency cells was not caused by the loss of NFE2L1 activity. Restoration of Ngly1 or AMPKα expression could significantly alleviate the energy metabolism defects.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0148.v1

SLC12A2: Based on lethality association and co-evolution analysis, a conserved Na/K/Cl ion transporter Ncc69 (human NKCC1/2, officially SLC12A1/2) was identified ubiquitously expressed and associated genetically with Ngly1 in Drosophila [90]. In Ncc69 or Ngly1 knockdown (KD) Drosophila, more than 30% of them showed severe seizures [91]. In mouse embryonic fibroblasts (MEFs), the homolog of Ncc69, NKCC1 was highly expressed and N-glycosylated, which was important for the function [92]. However, NGLY1 deficiency could disturbed the N-glycosylation of NKCC1 which was detrimental to the functionality of the protein. In mammals, NKCC1 was expressed specifically in secretory epithelia, such as salivary, sweat, and lacrimal glands, to promote the basolateral ion absorption and subsequent secretion [93]. The decrease of NKCC1 functionality might account for the alacrima and reduced production of saliva and sweat observed in patients with NGLY1 deficiency.

The variety of functions carried out by this enzyme may explain the diversity and varying severity of symptoms caused by mutations in this gene.

Mutations for NGLY1-related congenital disorders

In 2012, the first patient harboring pathogenic mutations in NGLY1 was identified by whole exome sequencing (WES) [3]. The bi-allelic mutations of NGLY1 could lead to a rare autosomal recessive congenital disorder of deglycosylation (CDD, OMIM#615273) which was characterized with global developmental delay, intellectual disability, microcephaly, hypotonia, motor function deficits, and poor or absent tear production [4; 20]. Other common features include seizures, optic atrophy, retinal pigmentary changes, cone dystrophy, delayed bone age, and joint hypermobility [94; 95]. It's revealed that this disorder mainly affects young children [3; 4]. Till now, 60 patients carrying bi-allelic NGLY1 mutations have been reported in literatures (n=54) and DECIPHER database (n=6) (Table 1). 43 different mutations were identified in these patients (Figure 6-A), with c.1201A>T (p.Arg401Ter) as the commonest (20%, 24/120) (Figure 6-A, -B). It's worth noting that nearly all missense mutations were located in the central transglutaminase-like core domain, and stop gain or frameshift in the C-terminal PAW domain (Figure 6-C). In these patients, stop gains accounted for 44.17% (53/120), missense 25.00% (30/120), frameshift 19.17% (23/120), splice donor 6.67% (8/120), splice acceptor 3.33% (4/120) and in-frame deletion 1.67 (2/120) (Figure 6-C). Nearly all NGLY1 mutations studied thus far are characterized by reduced NGLY1

protein levels and enzymatic activity [96].

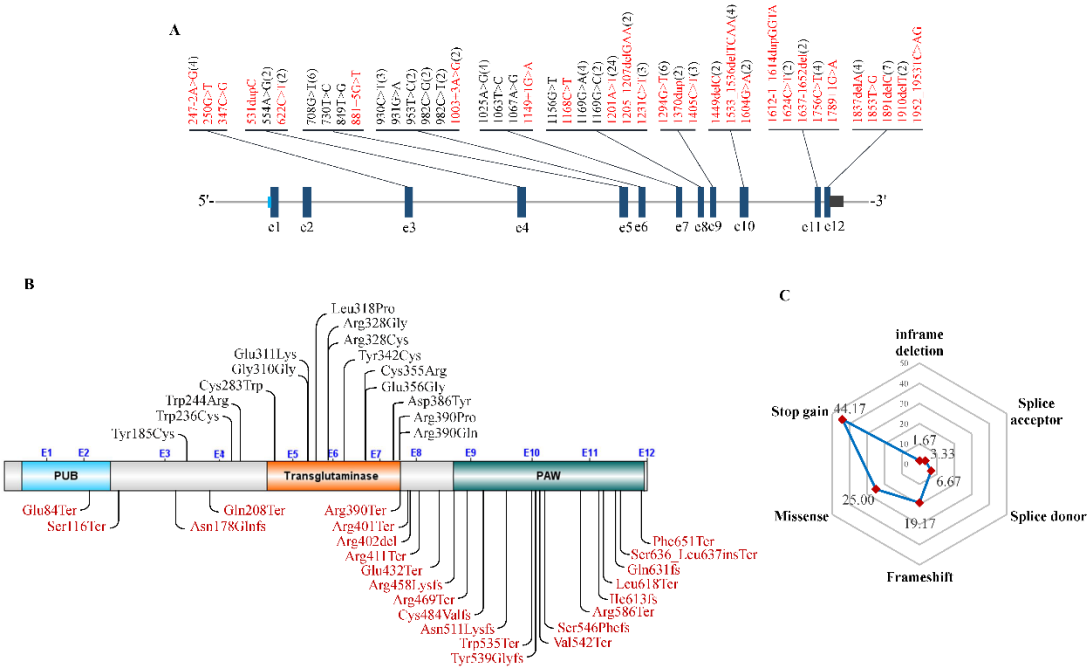

Figure 6. Characterization of mutations of NGLY1. A) mRNA locations; B) Amino acid locations; C) Different types of mutations.

For the missense mutations, missense3D was used to predict the structural changes introduced by an amino acid substitution based on the AlphFold predicted NGLY1 protein structure [27; 97]. Except for c.554A>G (p.Tyr185Cys) and c.1067A>G (p.Glu356Gly), other missense mutations possessed potential to cause structural damage to NGLY1, such as triggering clash alter (n=1, p.Cys283Trp), altering secondary structure (n=1, p.Arg390Pro), introducing buried Proline (n=2, p.Leu318Pro and p.Arg390Pro), breaking buried salt bridge (n=2, p.Arg328Cys and p.Arg328Gly), introducing buried charge switch (n=2, p.Glu311Lys and p.Leu318Pro), replacing buried charge residues (n=3, p.Arg328Cys, p.Arg328Gly and p.Arg390Gln), switching buried or exposed state (n=3, p.Cys283Trp, p.Cys355Arg and p.Arg390Pro), breaking buried H-bonds (n=4, p.Arg328Cys, p.Arg328Gly, p.Tyr342Cys and p.Arg390Pro), and expanding cavity volume (n=7, p.Trp236Cys, p.Trp244Arg, p.Leu318Pro, p.Arg328Cys, p.Tyr342Cys, p.Asp386Tyr and p.Arg390Pro) (Table 2). For p.Tyr185Cys (c.554A>G), it was identified homozygously in a 3-year-10-month old female (264084) recruited in the DECIPHER database. The phenotypes of this sample include moderate intellectual disability, delayed speech and language development, and echolalia, which was often observed in patient with NGLY1 deficiency [3; 20]. It's predicted by MutationTaster

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0148.v1

that this variant could generate an intraexonic splice donor site (GTCT|**g**tga, score = 0.76) resulting in an in-frame loss of 105 base pairs in exon 4. For c.1067A>G (p.Glu356Gly), it was classed PM2 (strong) + PP3 (supporting) + PP5 (supporting) and annotated as "Likely pathogenic" according to ACMG classification. The patient also carried a loss-of-function mutation in NGLY1 (c.1201A>T, p.Arg401Ter). Although predicted neutral to the protein structure of NGLY1, no NGLY1 proteins were detected in the muscle and skin fibroblasts of the patient by Western blot [22]. This indicated that the variant p.Glu356Gly could result in protein instability or destruction and be regarded as disease-causing mutation. Interestingly, nearly absent expressions of NGLY1 mRNAs were detected in fibroblast lines obtained from four patients with bi-allelic NGLY1 mutations [patient 1 (c.1205_1207delGAA, p.Arg402del and c.1624C>T, p.Arg542Ter); patient 2 (homozygous for c.1201A>T, p.Arg401Ter); patient 3 (c.931G>A, p.Glu311Lys and c.730T>C, p.Trp244Arg); patient 4 (c.622C>T, p.Gln208Ter and c.930C>T, p.Gly310=)] [88]. Although two missense mutations in patient 3 were predicted to affect the proper function of NGLY1, they actually impair the stability of messenger RNAs. In patient 4, c.930C>T was a synonymous mutation (p.Gly310=), but introduced a novel splice donor site. In the lymphoblastoid cell lines from a patient with compound NGLY missense mutations (c.953T>C, p.Leu318Pro and c.1169G>C, p.Arg390Pro), no proteins of NGLY1 were detected (data not shown). Therefore, it seemed quite likely that NGLY1 deficiency was resulted not only from nonsense-mediated mRNA decay (NMD) by nonsense or frameshift mutations [98], but also from mRNA or protein instability by missense mutations. This will be verified by in vitro and in vivo experiments soon.

Potential treatments for NGLY1-CDDG

With the rapid increase of research about NGLY1's structure, function and cellular pathways for cells and animal models in different species, several molecular methods were tested for the potential to treat NGLY1-deficient diseases.

Exogenous restoration of NGLY1 expression

Currently, double-knockout of NGLY1 have been used to study the enzymatic ability and molecular functions of the gene in several model organisms, such as yeast,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0148.v1

Drosophila, worms, mouse and rat. After comparing the spectrum of clinical phenotypes and histological analysis, Ngly1$^{-/-}$ rats displayed very similar features with human NGLY1-deficiency patients [72]. Although Ngly1 was ubiquitously expressed, and highly transcribed in brain, lung, heart, kidney, liver, placenta and testis [99], the neurological symptoms in most patients implied CNS be the most seriously affected organ [95].

AAV9s carrying a copy of human NGLY1 gene (AAV9-hNGLY1) were injected to Ngly1-KO SD rats during the weaning period via intracerebroventricular (i.c.v.) administration in order to reverse the deterioration of neuronal symptoms [100]. The re-expression of NGLY1 were identified in pons, thalamus, hippocampus, cerebral cortex and cerebellar Purkinje cells, but not detected in liver. The enzymatic activity of NGYL1 in Ngly1-/- rats was restored to a comparable level with that of wildtype controls. AAV9-hNGLY1 was safe and not leading to liver toxicity. The motor dysfunction caused by NGLY1 deficiency such as gait abnormalities, motor coordination and balance, were also significantly restored. However, the stride length and grip strength of limbs were not significantly ameliorated, this might be related to no expression of functional NGLY1 in muscles. This was the first time to restore the expression of NGLY1 in NGLY1-deficient mammalian model. However, there are many questions to be answered: 1) if rat was suitable for study the underlying molecular mechanisms of NGLY1-CDDG to explore potential drug targets. 2) Since NGLY1 was ubiquitously expressed, whether CNS was the most suitable organ to restore the exogenous NGLY1. If the injection of AVV9-NGLY1 in more than one organ could ameliorate the clinical phenotypes to a higher degree. 3) Since primates were most closely related to human in evolution, whether primates with NGLY1 deficiency might be the most appropriate animal model to study the pathogenesis of NGLY1-CDDG and to develop curable drugs.

Targeting ENGase with small inhibitors

It had been reported that endo-β-N-acetylglucosaminidase (ENGASE) could cleave the glycan moieties from N-linked glycoproteins at the beta-N-acetylglucosaminide, releasing a free GlcNAc and proteins with a single GlcNAc residue linked to the Asn residual (N-GlcNAc proteins) (Figure 4) [55]. The N-GlcNAc proteins are prone to aggregate, resulting in the dysfunction of ERAD process. Addition deletion of

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0148.v1

ENGASE could partially rescue of the lethality and alleviate the phenotypes of the Ngly1-loss mice [101]. Therefore, inhibition of ENGASE might be a potential therapeutic site for NGLY1-deficient diseases. Currently, several novel ENGASE inhibitors have been discovered, including Proton Pump Inhibitors (PPIs), Lansoprazole, Rabeprazole and Omeprazole [102]. PPIs could immediately be considered therapeutic drugs for treating NGLY1-deficiency because of its existing safety and pharmacokinetics profile.

Inhibition FOXB6 (Fbs2)

FOXB6 (also known as Fbs2) is a component of the SCF (SKP1-cullin-F-box) complex and functions as a substrate-binding adaptor [62]. In NGLY1 deficient cells, FOXB6 was overexpressed and resulted in cytotoxic impairment of the proteasome activity. The overexpressed FOXB6 not only induced ubiquitination of NFE2L1, but also inhibited the processing by DDI2 and nuclear localization in cells without NGLY1. Interestingly, addition knockout of Foxb6 could make Ngly1-deficient mice ($Foxb6^{-/-}$;$Ngly1^{-/-}$) viable and exhibiting normal motor functions [103]. The survival ratio at P0 was 2 times higher than in Engase; Ngly1 dKO mice [101]. These indicate that FOXB6 might be a promising drug target to treatment NGLY1-CDDG.

Activation of NFE2L2

It had been commonly accepted that inactivation of NFE2F1 was an important factor for the pathogenesis of NGLY1-CDDG. Functional loss of NFE2F1 could result in inhibition of proteasome function, mitochondrial dysfunction and immune dysregulation in NGLY1-deficient cells under proteotoxic stress [104]. In mammals, as a close homologue of NFE2F1, NFE2L2 (NRF2) was responsible for the regulation the expression of similar proteasome subunit genes as NFE2L1 did under oxidative stresses [105], and also the expression of genes for autophagy and mitophagy [106; 107]. Not like the ER-bound NFE2FL1, NFE2L2 was cytosolic and non-N-glycosylated without relying on NGLY1 for its activation [108]. Increased expression of NFE2F2 could promote mitophagy and rescue the mitochondrial and immune homeostasis in $Ngly1^{-/-}$ cells. It had been verified that KEAP1 could strongly bind with and mediate the ubiquitination and degradation of NFE2L2. Therefore, chemical inhibitors to disrupt the KEAP1-NFE2L2 interaction might be a promising drug target. Interestingly, a natural inhibitor for KEAP1 derived from cruciferous vegetables (such as broccoli, cauliflower, kale,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0148.v1

kohlrabi, brussel sprouts and cabbage), Sulforaphane could efficiently binding with KEAP1 and robustly increase NFE2L2 protein level to promote the expression of proteasome subunit genes and mitophagy-related genes in NGLY1-deficient cells [104]. The disrupted mitochondria and dysregulated immune response to mtDNAs were significantly decreased after administration of Sulforaphane, which indicated that KEAP1-NFE2L2 axis was a novel therapeutic site for correcting the abnormalities of NGLY1 diseases.

Perspectives

Research associated with NGLY1 physiology, structure, function and diseases at molecular, cellular and model animal levels have been conducted for almost 50 years since 1977. However, its alternative splicing patterns, cellular expression and localization were not clear. Since ER stress was not directly caused by loss of NGLY1 in many models, it is desired to identify the direct targets of NGLY1 and their contribution to the pathogenesis of NGLY1 deficiency. For locations and types of the NGLY1 mutation for CDDG, almost all missense mutations were in the central transglutaminase domain, which might affect the de-glycosylating enzyme activity of NGLY1. However, a great portion of the missense mutations could result in the loss of mRNA or protein detected by quantification PCR or Western blot. The underlying mechanisms are still not revealed. It had been reported that many cellular pathways could be disturbed by the loss of NGLY1, such as NFE2L1 pathway, Creb1/Atf1-AQP pathway, BMP pathway, AMPK pathway, SLC12A2 ion transporter and so on. These might be the underlying reasons for a constellation of clinical phenotypes of NGLY1 deficiency. Since many pathways function tissue specifically and at different development stages, these make the treatment much more challenging.

Supplemental data

Supplemental data include three figures and three tables.

Declaration of interests

There were no competing interests among the authors and fundings.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0148.v1

Acknowledgments

Key Research and Development Project of Deyang City's Science and Technology Bureau (21ZDYF0011 to GL.), Special Fund for Incubation Projects of Deyang People's Hospital (FHG202004 to GL.).

Author's contributions

G.L and C.H conceived the project, wrote and revised the manuscript. M.X and J.W collected data and recruited mutations. G.L and C.H were corresponding authors.

Data and code availability

All other data are available on request.

References

[1] J. Hwang, and L. Qi, Quality Control in the Endoplasmic Reticulum: Crosstalk between ERAD and UPR pathways. Trends in biochemical sciences 43 (2018) 593-605.
[2] T. Suzuki, H. Park, and W.J. Lennarz, Cytoplasmic peptide:N-glycanase (PNGase) in eukaryotic cells: occurrence, primary structure, and potential functions. FASEB J 16 (2002) 635-41.
[3] A.C. Need, V. Shashi, Y. Hitomi, K. Schoch, K.V. Shianna, M.T. McDonald, M.H. Meisler, and D.B. Goldstein, Clinical application of exome sequencing in undiagnosed genetic conditions. Journal of medical genetics 49 (2012) 353-61.
[4] G.M. Enns, V. Shashi, M. Bainbridge, M.J. Gambello, F.R. Zahir, T. Bast, R. Crimian, K. Schoch, J. Platt, R. Cox, J.A. Bernstein, M. Scavina, R.S. Walter, A. Bibb, M. Jones, M. Hegde, B.H. Graham, A.C. Need, A. Oviedo, C.P. Schaaf, S. Boyle, A.J. Butte, R. Chen, R. Chen, M.J. Clark, R. Haraksingh, F.C. Consortium, T.M. Cowan, P. He, S. Langlois, H.Y. Zoghbi, M. Snyder, R.A. Gibbs, H.H. Freeze, and D.B. Goldstein, Mutations in NGLY1 cause an inherited disorder of the endoplasmic reticulum-associated degradation pathway. Genet Med 16 (2014) 751-8.
[5] N. Takahashi, Demonstration of a new amidase acting on glycopeptides. Biochemical and biophysical research communications 76 (1977) 1194-201.
[6] K. Sugiyama, H. Ishihara, S. Tejima, and N. Takahashi, Demonstration of a new glycopeptidase, from jack-bean meal, acting on aspartylglucosylamine linkages. Biochemical and biophysical research communications 112 (1983) 155-60.
[7] T.H. Plummer, Jr., A.W. Phelan, and A.L. Tarentino, Detection and quantification of peptide-N4-(N-acetyl-beta-glucosaminyl)asparagine amidases. European journal of biochemistry 163 (1987) 167-73.
[8] A. Diepold, G. Li, W.J. Lennarz, T. Nurnberger, and F. Brunner, The Arabidopsis AtPNG1 gene encodes a peptide: N-glycanase. The Plant journal : for cell and molecular biology 52 (2007) 94-104.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0148.v1

[9] N. Takahashi, and H. Nishibe, Some characteristics of a new glycopeptidase acting on aspartylglycosylamine linkages. Journal of biochemistry 84 (1978) 1467-73.
[10] T.H. Plummer, Jr., and A.L. Tarentino, Facile cleavage of complex oligosaccharides from glycopeptides by almond emulsin peptide: N-glycosidase. The Journal of biological chemistry 256 (1981) 10243-6.
[11] H. Nishibe, and N. Takahashi, The release of carbohydrate moieties from human fibrinogen by almond glycopeptidase without alteration in fibrinogen clottability. Biochimica et biophysica acta 661 (1981) 274-9.
[12] A.L. Tarentino, and T.H. Plummer, Jr., Oligosaccharide accessibility to peptide:N-glycosidase as promoted by protein-unfolding reagents. The Journal of biological chemistry 257 (1982) 10776-80.
[13] T.H. Plummer, Jr., J.H. Elder, S. Alexander, A.W. Phelan, and A.L. Tarentino, Demonstration of peptide:N-glycosidase F activity in endo-beta-N-acetylglucosaminidase F preparations. The Journal of biological chemistry 259 (1984) 10700-4.
[14] K. Ishii, M. Iwasaki, S. Inoue, P.T. Kenny, H. Komura, and Y. Inoue, Free sialooligosaccharides found in the unfertilized eggs of a freshwater trout, Plecoglossus altivelis. A large storage pool of complex-type bi-, tri-, and tetraantennary sialooligosaccharides. The Journal of biological chemistry 264 (1989) 1623-30.
[15] A. Seko, K. Kitajima, M. Iwasaki, S. Inoue, and Y. Inoue, Structural studies of fertilization-associated carbohydrate-rich glycoproteins (hyosophorin) isolated from the fertilized and unfertilized eggs of flounder, Paralichthys olivaceus. Presence of a novel penta-antennary N-linked glycan chain in the tandem repeating glycopeptide unit of hyosophorin. The Journal of biological chemistry 264 (1989) 15922-9.
[16] A. Seko, K. Kitajima, Y. Inoue, and S. Inoue, Peptide:N-glycosidase activity found in the early embryos of Oryzias latipes (Medaka fish). The first demonstration of the occurrence of peptide:N-glycosidase in animal cells and its implication for the presence of a de-N-glycosylation system in living organisms. The Journal of biological chemistry 266 (1991) 22110-4.
[17] T. Suzuki, A. Seko, K. Kitajima, Y. Inoue, and S. Inoue, Identification of peptide:N-glycanase activity in mammalian-derived cultured cells. Biochemical and biophysical research communications 194 (1993) 1124-30.
[18] N. Ftouhi-Paquin, C.R. Hauer, R.F. Stack, A.L. Tarentino, and T.H. Plummer, Jr., Molecular cloning, primary structure, and properties of a new glycoamidase from the fungus Aspergillus tubigensis. The Journal of biological chemistry 272 (1997) 22960-5.
[19] T. Suzuki, H. Park, K. Kitajima, and W.J. Lennarz, Peptides glycosylated in the endoplasmic reticulum of yeast are subsequently deglycosylated by a soluble peptide: N-glycanase activity. The Journal of biological chemistry 273 (1998) 21526-30.
[20] C. Lam, C. Ferreira, D. Krasnewich, C. Toro, L. Latham, W.M. Zein, T. Lehky, C. Brewer, E.H. Baker, A. Thurm, C.A. Farmer, S.D. Rosenzweig, J.J. Lyons, J.M. Schreiber, A. Gropman, S. Lingala, M.G. Ghany, B. Solomon, E. Macnamara, M. Davids, C.A. Stratakis, V. Kimonis, W.A. Gahl, and L. Wolfe, Prospective phenotyping of NGLY1-CDDG, the first congenital disorder of deglycosylation. Genet Med 19 (2017) 160-168.
[21] H.A. Haijes, M.G.M. de Sain-van der Velden, H. Prinsen, A.P. Willems, M. van der Ham, J. Gerrits, M.H. Couse, J.M. Friedman, C.D.M. van Karnebeek, K.A. Selby, P.M. van Hasselt, N.M. Verhoeven-Duif, and J.J.M. Jans,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0148.v1

Aspartylglycosamine is a biomarker for NGLY1-CDDG, a congenital disorder of deglycosylation. Mol Genet Metab 127 (2019) 368-372.

[22] D.M. Panneman, S.B. Wortmann, C.A. Haaxma, P.M. van Hasselt, N.I. Wolf, Y. Hendriks, B. Kusters, S. van Emst-de Vries, E. van de Westerlo, W.J.H. Koopman, L. Wintjes, F. van den Brandt, M. de Vries, D.J. Lefeber, J.A.M. Smeitink, and R.J. Rodenburg, Variants in NGLY1 lead to intellectual disability, myoclonus epilepsy, sensorimotor axonal polyneuropathy and mitochondrial dysfunction. Clin Genet 97 (2020) 556-566.

[23] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Zidek, A. Potapenko, A. Bridgland, C. Meyer, S.A.A. Kohl, A.J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A.W. Senior, K. Kavukcuoglu, P. Kohli, and D. Hassabis, Highly accurate protein structure prediction with AlphaFold. Nature 596 (2021) 583-589.

[24] B. Meszaros, G. Erdos, and Z. Dosztanyi, IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. Nucleic Acids Res 46 (2018) W329-W337.

[25] T. Suzuki, H. Park, E.A. Till, and W.J. Lennarz, The PUB domain: a putative protein-protein interaction domain implicated in the ubiquitin-proteasome pathway. Biochemical and biophysical research communications 287 (2001) 1083-7.

[26] T. Doerks, R.R. Copley, J. Schultz, C.P. Ponting, and P. Bork, Systematic identification of novel protein domain families associated with nuclear functions. Genome Res 12 (2002) 47-56.

[27] M. Varadi, S. Anyango, M. Deshpande, S. Nair, C. Natassia, G. Yordanova, D. Yuan, O. Stroe, G. Wood, A. Laydon, A. Zidek, T. Green, K. Tunyasuvunakool, S. Petersen, J. Jumper, E. Clancy, R. Green, A. Vora, M. Lutfi, M. Figurnov, A. Cowie, N. Hobbs, P. Kohli, G. Kleywegt, E. Birney, D. Hassabis, and S. Velankar, AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic Acids Res 50 (2022) D439-D444.

[28] K. Hofmann, and P. Bucher, The UBA domain: a sequence motif present in multiple enzyme classes of the ubiquitination pathway. Trends in biochemical sciences 21 (1996) 172-3.

[29] M.D. Allen, A. Buchberger, and M. Bycroft, The PUB domain functions as a p97 binding module in human peptide N-glycanase. The Journal of biological chemistry 281 (2006) 25502-8.

[30] Y. Kamiya, Y. Uekusa, A. Sumiyoshi, H. Sasakawa, T. Hirao, T. Suzuki, and K. Kato, NMR characterization of the interaction between the PUB domain of peptide:N-glycanase and ubiquitin-like domain of HR23. FEBS Lett 586 (2012) 1141-6.

[31] V. Schaeffer, M. Akutsu, M.H. Olma, L.C. Gomes, M. Kawasaki, and I. Dikic, Binding of OTULIN to the PUB domain of HOIP controls NF-kappaB signaling. Mol Cell 54 (2014) 349-61.

[32] J.C. Christianson, J.A. Olzmann, T.A. Shaler, M.E. Sowa, E.J. Bennett, C.M. Richter, R.E. Tyler, E.J. Greenblatt, J.W. Harper, and R.R. Kopito, Defining human ERAD networks through an integrative mapping strategy. Nature cell biology 14 (2011) 93-105.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0148.v1

[33] T. Suzuki, K. Kitajima, Y. Emori, Y. Inoue, and S. Inoue, Site-specific de-N-glycosylation of diglycosylated ovalbumin in hen oviduct by endogenous peptide: N-glycanase as a quality control system for newly synthesized proteins. Proceedings of the National Academy of Sciences of the United States of America 94 (1997) 6244-9.

[34] S. Katiyar, S. Joshi, and W.J. Lennarz, The retrotranslocation protein Derlin-1 binds peptide:N-glycanase to the endoplasmic reticulum. Molecular biology of the cell 16 (2005) 4584-94.

[35] J.H. Lee, J.M. Choi, C. Lee, K.J. Yi, and Y. Cho, Structure of a peptide:N-glycanase-Rad23 complex: insight into the deglycosylation for denatured glycoproteins. Proceedings of the National Academy of Sciences of the United States of America 102 (2005) 9144-9.

[36] G. Zhao, X. Zhou, L. Wang, G. Li, C. Kisker, W.J. Lennarz, and H. Schindelin, Structure of the mouse peptide N-glycanase-HR23 complex suggests co-evolution of the endoplasmic reticulum-associated degradation and DNA repair pathways. The Journal of biological chemistry 281 (2006) 13751-13761.

[37] E.M. Taga, A. Waheed, and R.L. Van Etten, Structural and chemical characterization of a homogeneous peptide N-glycosidase from almond. Biochemistry 23 (1984) 815-22.

[38] S. Maerz, Y. Funakoshi, Y. Negishi, T. Suzuki, and S. Seiler, The Neurospora peptide:N-glycanase ortholog PNG1 is essential for cell polarity despite its lack of enzymatic activity. The Journal of biological chemistry 285 (2010) 2326-32.

[39] L. Chen, and K. Madura, Rad23 promotes the targeting of proteolytic substrates to the proteasome. Mol Cell Biol 22 (2002) 4902-13.

[40] D. Lambertson, L. Chen, and K. Madura, Investigating the importance of proteasome-interaction for Rad23 function. Curr Genet 42 (2003) 199-208.

[41] N.P. Dantuma, C. Heinen, and D. Hoogstraten, The ubiquitin receptor Rad23: at the crossroads of nucleotide excision repair and proteasomal degradation. DNA Repair (Amst) 8 (2009) 449-60.

[42] T. van Eeuwen, Y. Shim, H.J. Kim, T. Zhao, S. Basu, B.A. Garcia, C.D. Kaplan, J.H. Min, and K. Murakami, Cryo-EM structure of TFIIH/Rad4-Rad23-Rad33 in damaged DNA opening in nucleotide excision repair. Nat Commun 12 (2021) 3338.

[43] X. Zhou, G. Zhao, J.J. Truglio, L. Wang, G. Li, W.J. Lennarz, and H. Schindelin, Structural and biochemical studies of the C-terminal domain of mouse peptide-N-glycanase identify it as a mannose-binding module. Proceedings of the National Academy of Sciences of the United States of America 103 (2006) 17214-9.

[44] K. Takahashi, and M. Nei, Efficiencies of fast algorithms of phylogenetic inference under the criteria of maximum parsimony, minimum evolution, and maximum likelihood when a large number of sequences are used. Mol Biol Evol 17 (2000) 1251-8.

[45] Y. Rao, and V. Haucke, Membrane shaping by the Bin/amphiphysin/Rvs (BAR) domain protein superfamily. Cell Mol Life Sci 68 (2011) 3983-93.

[46] M. Masuda, and N. Mochizuki, Structural characteristics of BAR domain superfamily to sculpt the membrane. Semin Cell Dev Biol 21 (2010) 391-8.

[47] H. Zhao, A. Michelot, E.V. Koskela, V. Tkach, D. Stamou, D.G. Drubin, and P. Lappalainen, Membrane-sculpting BAR domains generate stable lipid microdomains. Cell Rep 4 (2013) 1213-23.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0148.v1

[48] T. Nishimura, N. Morone, and S. Suetsugu, Membrane re-modelling by BAR domain superfamily proteins via molecular and non-molecular factors. Biochem Soc Trans 46 (2018) 379-389.
[49] T. Suzuki, K. Tanabe, I. Hara, N. Taniguchi, and A. Colavita, Dual enzymatic properties of the cytoplasmic peptide: N-glycanase in C. elegans. Biochemical and biophysical research communications 358 (2007) 837-41.
[50] T. Kato, A. Kawahara, H. Ashida, and K. Yamamoto, Unique peptide:N-glycanase of Caenorhabditis elegans has activity of protein disulphide reductase as well as of deglycosylation. Journal of biochemistry 142 (2007) 175-81.
[51] Y. Masahara-Negishi, A. Hosomi, M. Della Mea, D. Serafini-Fracassini, and T. Suzuki, A plant peptide: N-glycanase orthologue facilitates glycoprotein ER-associated degradation in yeast. Biochimica et biophysica acta 1820 (2012) 1457-62.
[52] H.J. Joshi, Y. Narimatsu, K.T. Schjoldager, H.L.P. Tytgat, M. Aebi, H. Clausen, and A. Halim, SnapShot: O-Glycosylation Pathways across Kingdoms. Cell 172 (2018) 632-632 e2.
[53] C.Y. Chung, N.I. Majewska, Q. Wang, J.T. Paul, and M.J. Betenbaugh, SnapShot: N-Glycosylation Processing Pathways across Kingdoms. Cell 171 (2017) 258-258 e1.
[54] F.U. Hartl, Protein Misfolding Diseases. Annu Rev Biochem 86 (2017) 21-26.
[55] C. Huang, Y. Harada, A. Hosomi, Y. Masahara-Negishi, J. Seino, H. Fujihira, Y. Funakoshi, T. Suzuki, N. Dohmae, and T. Suzuki, Endo-beta-N-acetylglucosaminidase forms N-GlcNAc protein aggregates during ER-associated degradation in Ngly1-defective cells. Proceedings of the National Academy of Sciences of the United States of America 112 (2015) 1398-403.
[56] F. Chiti, and C.M. Dobson, Protein Misfolding, Amyloid Formation, and Human Disease: A Summary of Progress Over the Last Decade. Annu Rev Biochem 86 (2017) 27-68.
[57] B.P. Phillips, N. Gomez-Navarro, and E.A. Miller, Protein quality control in the endoplasmic reticulum. Curr Opin Cell Biol 65 (2020) 96-102.
[58] H. Kim, A. Bhattacharya, and L. Qi, Endoplasmic reticulum quality control in cancer: Friend or foe. Semin Cancer Biol 33 (2015) 25-33.
[59] C. Schuberth, and A. Buchberger, Membrane-bound Ubx2 recruits Cdc48 to ubiquitin ligases and their substrates to ensure efficient ER-associated protein degradation. Nature cell biology 7 (2005) 999-1006.
[60] E. Jarosch, C. Taxis, C. Volkwein, J. Bordallo, D. Finley, D.H. Wolf, and T. Sommer, Protein dislocation from the ER requires polyubiquitination and the AAA-ATPase Cdc48. Nature cell biology 4 (2002) 134-9.
[61] Y. Yoshida, T. Chiba, F. Tokunaga, H. Kawasaki, K. Iwai, T. Suzuki, Y. Ito, K. Matsuoka, M. Yoshida, K. Tanaka, and T. Tai, E3 ubiquitin ligase that recognizes sugar chains. Nature 418 (2002) 438-42.
[62] Y. Yoshida, F. Tokunaga, T. Chiba, K. Iwai, K. Tanaka, and T. Tai, Fbs2 is a new member of the E3 ubiquitin ligase family that recognizes sugar chains. The Journal of biological chemistry 278 (2003) 43877-84.
[63] Y. Yoshida, E. Adachi, K. Fukiya, K. Iwai, and K. Tanaka, Glycoprotein-specific ubiquitin ligases recognize N-glycans in unfolded substrates. EMBO Rep 6 (2005) 239-44.
[64] T. Mizushima, Y. Yoshida, T. Kumanomidou, Y. Hasegawa, A. Suzuki, T. Yamane, and K. Tanaka, Structural basis for the selection of glycosylated substrates by SCF(Fbs1) ubiquitin ligase. Proceedings of the National Academy of Sciences

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0148.v1

of the United States of America 104 (2007) 5777-81.
[65] Y. Yoshida, T. Mizushima, and K. Tanaka, Sugar-Recognizing Ubiquitin Ligases: Action Mechanisms and Physiology. Front Physiol 10 (2019) 104.
[66] L. Qi, B. Tsai, and P. Arvan, New Insights into the Physiological Role of Endoplasmic Reticulum-Associated Degradation. Trends Cell Biol 27 (2017) 430-440.
[67] J.C. Maynard, H. Fujihira, G.E. Dolgonos, T. Suzuki, and A.L. Burlingame, Cytosolic N-GlcNAc proteins are formed by the action of endo-beta-N-acetylglucosaminidase. Biochemical and biophysical research communications 530 (2020) 719-724.
[68] N. Sovolyova, S. Healy, A. Samali, and S.E. Logue, Stressed to death - mechanisms of ER stress-induced cell death. Biol Chem 395 (2014) 1-13.
[69] E. Kario, B. Tirosh, H.L. Ploegh, and A. Navon, N-linked glycosylation does not impair proteasomal degradation but affects class I major histocompatibility complex presentation. The Journal of biological chemistry 283 (2008) 244-254.
[70] D. Blom, C. Hirsch, P. Stern, D. Tortorella, and H.L. Ploegh, A glycosylated type I membrane protein becomes cytosolic when peptide: N-glycanase is compromised. EMBO J 23 (2004) 650-8.
[71] K.G. Owings, J.B. Lowry, Y. Bi, M. Might, and C.Y. Chow, Transcriptome and functional analysis in a Drosophila model of NGLY1 deficiency provides insight into therapeutic approaches. Hum Mol Genet 27 (2018) 1055-1066.
[72] M. Asahina, R. Fujinawa, S. Nakamura, K. Yokoyama, R. Tozawa, and T. Suzuki, Ngly1 -/- rats develop neurodegenerative phenotypes and pathological abnormalities in their peripheral and central nervous systems. Hum Mol Genet 29 (2020) 1635-1647.
[73] M.A. Tambe, B.G. Ng, and H.H. Freeze, N-Glycanase 1 Transcriptionally Regulates Aquaporins Independent of Its Enzymatic Activity. Cell Rep 29 (2019) 4620-4631 e4.
[74] W.F. Mueller, P. Jakob, H. Sun, S. Clauder-Munster, S. Ghidelli-Disse, D. Ordonez, M. Boesche, M. Bantscheff, P. Collier, B. Haase, V. Benes, M. Paulsen, P. Sehr, J. Lewis, G. Drewes, and L.M. Steinmetz, Loss of N-Glycanase 1 Alters Transcriptional and Translational Regulation in K562 Cell Lines. G3 (Bethesda) 10 (2020) 1585-1597.
[75] H.M. Kim, J.W. Han, and J.Y. Chan, Nuclear Factor Erythroid-2 Like 1 (NFE2L1): Structure, function and regulation. Gene 584 (2016) 17-25.
[76] J. Steffen, M. Seeger, A. Koch, and E. Kruger, Proteasomal degradation is transcriptionally controlled by TCF11 via an ERAD-dependent feedback loop. Mol Cell 40 (2010) 147-58.
[77] N.J. Lehrbach, and G. Ruvkun, Proteasome dysfunction triggers activation of SKN-1A/Nrf1 by the aspartic protease DDI-1. Elife 5 (2016).
[78] S. Koizumi, T. Irie, S. Hirayama, Y. Sakurai, H. Yashiroda, I. Naguro, H. Ichijo, J. Hamazaki, and S. Murata, The aspartyl protease DDI2 activates Nrf1 to compensate for proteasome dysfunction. Elife 5 (2016).
[79] P. Murphy, and A. Kolsto, Expression of the bZIP transcription factor TCF11 and its potential dimerization partners during development. Mech Dev 97 (2000) 141-8.
[80] C. Delporte, A. Bryla, and J. Perret, Aquaporins in Salivary Glands: From Basic Research to Clinical Applications. Int J Mol Sci 17 (2016).
[81] H. Zhi, and W.T. Yuan, [Expression of aquaporin 3, 4, and 8 in colonic mucosa of rat models with slow transit constipation]. Zhonghua Wei Chang Wai Ke Za Zhi

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0148.v1

14 (2011) 459-61.
[82] A. Galeone, S.Y. Han, C. Huang, A. Hosomi, T. Suzuki, and H. Jafar-Nejad, Tissue-specific regulation of BMP signaling by Drosophila N-glycanase 1. Elife 6 (2017).
[83] A. Galeone, J.M. Adams, S. Matsuda, M.F. Presa, A. Pandey, S.Y. Han, Y. Tachida, H. Hirayama, T. Vaccari, T. Suzuki, C.M. Lutz, M. Affolter, A. Zuberi, and H. Jafar-Nejad, Regulation of BMP4/Dpp retrotranslocation and signaling by de-glycosylation. Elife 9 (2020).
[84] P. Bakrania, M. Efthymiou, J.C. Klein, A. Salt, D.J. Bunyan, A. Wyatt, C.P. Ponting, A. Martin, S. Williams, V. Lindley, J. Gilmore, M. Restori, A.G. Robson, M.M. Neveu, G.E. Holder, J.R. Collin, D.O. Robinson, P. Farndon, H. Johansen-Berg, D. Gerrelli, and N.K. Ragge, Mutations in BMP4 cause eye, brain, and digit developmental anomalies: overlap between the BMP4 and hedgehog signaling pathways. Am J Hum Genet 82 (2008) 304-19.
[85] T. Obara-Ishihara, J. Kuhlman, L. Niswander, and D. Herzlinger, The surface ectoderm is essential for nephric duct formation in intermediate mesoderm. Development 126 (1999) 1103-8.
[86] L. Tirosh-Finkel, A. Zeisel, M. Brodt-Ivenshitz, A. Shamai, Z. Yao, R. Seger, E. Domany, and E. Tzahor, BMP-mediated inhibition of FGF signaling promotes cardiomyocyte differentiation of anterior heart field progenitors. Development 137 (2010) 2989-3000.
[87] L.M. Reis, R.C. Tyler, K.F. Schilter, O. Abdul-Rahman, J.W. Innis, B.A. Kozel, A.S. Schneider, T.M. Bardakjian, E.J. Lose, D.M. Martin, U. Broeckel, and E.V. Semina, BMP4 loss-of-function mutations in developmental eye disorders including SHORT syndrome. Hum Genet 130 (2011) 495-504.
[88] J. Kong, M. Peng, J. Ostrovsky, Y.J. Kwon, O. Oretsky, E.M. McCormick, M. He, Y. Argon, and M.J. Falk, Mitochondrial function requires NGLY1. Mitochondrion 38 (2018) 6-16.
[89] S.Y. Han, A. Pandey, T. Moore, A. Galeone, L. Duraine, T.M. Cowan, and H. Jafar-Nejad, A conserved role for AMP-activated protein kinase in NGLY1 deficiency. PLoS Genet 16 (2020) e1009258.
[90] D.M. Talsness, K.G. Owings, E. Coelho, G. Mercenne, J.M. Pleinis, R. Partha, K.A. Hope, A.R. Zuberi, N.L. Clark, C.M. Lutz, A.R. Rodan, and C.Y. Chow, A Drosophila screen identifies NKCC1 as a modifier of NGLY1 deficiency. Elife 9 (2020).
[91] Z.M. Rusan, O.A. Kingsford, and M.A. Tanouye, Modeling glial contributions to seizures and epileptogenesis: cation-chloride cotransporters in Drosophila melanogaster. PLoS One 9 (2014) e101117.
[92] R. Singh, M.M. Almutairi, R. Pacheco-Andrade, M.Y. Almiahuob, and M. Di Fulvio, Impact of Hybrid and Complex N-Glycans on Cell Surface Targeting of the Endogenous Chloride Cotransporter Slc12a2. Int J Cell Biol 2015 (2015) 505294.
[93] E. Delpire, and K.B. Gagnon, Na(+) -K(+) -2Cl(-) Cotransporter (NKCC) Physiological Function in Nonpolarized Cells and Transporting Epithelia. Compr Physiol 8 (2018) 871-901.
[94] K. Abuduxikuer, L. Zou, L. Wang, L. Chen, and J.S. Wang, Novel NGLY1 gene variants in Chinese children with global developmental delay, microcephaly, hypotonia, hypertransaminasemia, alacrimia, and feeding difficulty. J Hum Genet 65 (2020) 387-396.
[95] A. Kariminejad, M. Shakiba, M. Shams, P. Namiranian, M. Eghbali, S. Talebi, M.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0148.v1

Makvand, J. Jaeken, H. Najmabadi, and R.C. Hennekam, NGLY1 deficiency: Novel variants and literature review. Eur J Med Genet 64 (2021) 104146.

[96] P. He, J.E. Grotzke, B.G. Ng, M. Gunel, H. Jafar-Nejad, P. Cresswell, G.M. Enns, and H.H. Freeze, A congenital disorder of deglycosylation: Biochemical characterization of N-glycanase 1 deficiency in patient fibroblasts. Glycobiology 25 (2015) 836-44.

[97] S. Ittisoponpisan, S.A. Islam, T. Khanna, E. Alhuzimi, A. David, and M.J.E. Sternberg, Can Predicted Protein 3D Structures Provide Reliable Insights into whether Missense Variants Are Disease Associated? J Mol Biol 431 (2019) 2197-2212.

[98] E.D. Karousis, and O. Muhlemann, Nonsense-Mediated mRNA Decay Begins Where Translation Ends. Cold Spring Harb Perspect Biol 11 (2019).

[99] T. Suzuki, M.A. Kwofie, and W.J. Lennarz, Ngly1, a mouse gene encoding a deglycosylating enzyme implicated in proteasomal degradation: expression, genomic organization, and chromosomal mapping. Biochemical and biophysical research communications 304 (2003) 326-32.

[100] M. Asahina, R. Fujinawa, H. Hirayama, R. Tozawa, Y. Kajii, and T. Suzuki, Reversibility of motor dysfunction in the rat model of NGLY1 deficiency. Mol Brain 14 (2021) 91.

[101] H. Fujihira, Y. Masahara-Negishi, M. Tamura, C. Huang, Y. Harada, S. Wakana, D. Takakura, N. Kawasaki, N. Taniguchi, G. Kondoh, T. Yamashita, Y. Funakoshi, and T. Suzuki, Lethality of mice bearing a knockout of the Ngly1-gene is partially rescued by the additional deletion of the Engase gene. PLoS Genet 13 (2017) e1006696.

[102] Y. Bi, M. Might, H. Vankayalapati, and B. Kuberan, Repurposing of Proton Pump Inhibitors as first identified small molecule inhibitors of endo-beta-N-acetylglucosaminidase (ENGase) for the treatment of NGLY1 deficiency, a rare genetic disease. Bioorg Med Chem Lett 27 (2017) 2962-2966.

[103] Y. Yoshida, M. Asahina, A. Murakami, J. Kawawaki, M. Yoshida, R. Fujinawa, K. Iwai, R. Tozawa, N. Matsuda, K. Tanaka, and T. Suzuki, Loss of peptide:N-glycanase causes proteasome dysfunction mediated by a sugar-recognizing ubiquitin ligase. Proceedings of the National Academy of Sciences of the United States of America 118 (2021).

[104] K. Yang, R. Huang, H. Fujihira, T. Suzuki, and N. Yan, N-glycanase NGLY1 regulates mitochondrial homeostasis and inflammation through NRF1. J Exp Med 215 (2018) 2600-2616.

[105] M.K. Kwak, N. Wakabayashi, J.L. Greenlaw, M. Yamamoto, and T.W. Kensler, Antioxidants enhance mammalian proteasome expression through the Keap1-Nrf2 signaling pathway. Mol Cell Biol 23 (2003) 8786-94.

[106] D.A. East, F. Fagiani, J. Crosby, N.D. Georgakopoulos, H. Bertrand, M. Schaap, A. Fowkes, G. Wells, and M. Campanella, PMI: a DeltaPsim independent pharmacological regulator of mitophagy. Chem Biol 21 (2014) 1585-96.

[107] M. Pajares, N. Jimenez-Moreno, A.J. Garcia-Yague, M. Escoll, M.L. de Ceballos, F. Van Leuven, A. Rabano, M. Yamamoto, A.I. Rojo, and A. Cuadrado, Transcription factor NFE2L2/NRF2 is a regulator of macroautophagy genes. Autophagy 12 (2016) 1902-1916.

[108] T.W. Kensler, N. Wakabayashi, and S. Biswal, Cell survival responses to environmental stresses via the Keap1-Nrf2-ARE pathway. Annu Rev Pharmacol Toxicol 47 (2007) 89-116.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0148.v1