Heliyon 10 (2024) e28787

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Functional prediction of the potential NGLY1 mutations associated with rare disease CDG

Shuying Yuan [a,1], Yanwen Chen [a,1], Lin Zou [b], Xinrong Lu [b], Ruijie Liu [a], Shaoxing Zhang [a], Yuxin Zhang [a], Cuiying Chen [c], Dongqing Cheng [a], Li Chen [b,**], Guiqin Sun [a,*]

[a] School of Medical Technology and Information Engineering, Zhejiang Chinese Medical University, Hangzhou, 310053, Zhejiang Province, China
[b] Department of Medical Microbiology and Parasitology, Key Laboratory of Medical Molecular Virology of Ministries of Education and Health, School of Basic Medical Sciences, Fudan University, Shanghai, 200032, China
[c] Department of Research and Development, SysDiagno Biotech, Nanjing, 211800, Jiangsu Province, China

ARTICLE INFO

Keywords:
Deglycosylation
Enzymatic activity
Gain of function mutation
PNGase
Rare disease

ABSTRACT

Genetic diseases are currently diagnosed by functional mutations. However, only some mutations are associated with disease. It is necessary to establish a quick prediction model for clinical screening. Pathogenic mutations in NGLY1 cause a rare autosomal recessive disease known as congenital disorder of deglycosylation (NGLY1-CDDG). Although NGLY1-CDDG can be diagnosed through gene sequencing, clinical relevance of a detected mutation in NGLY1 needs to be further confirmed. In this study, taken NGLY1-CDDG as an example, a comprehensive and practical predictive model for pathogenic mutations on NGLY1 through an NGLY1/Glycopeptide complex model was constructed, the binding sites of NGLY1 and glycopeptides were simulated, and an in vitro enzymatic assay system was established to facilitate quick clinical decisions for NGLY1-CDDG patients. The docking model covers 42 % of reported NGLY1-CDDG missense mutations (5/12). All reported mutations were subjected to in vitro enzymatic assay in which 18 mutations were dysfunctional (18/30). In addition, a full spectrum of functional R328 mutations was assayed and 11 mutations were dysfunctional (11/19). In this study, a model of NGLY1 and glycopeptides was built for potential functional mutations in NGLY1. In addition, the effect of potential regulatory compounds, including N-acetyl-L-cysteine and dithiothreitol, on NGLY1 was examined. The established in vitro assay may serve as a standard protocol to facilitate rapid diagnosis of all mutations in NGLY1-CDDG. This method could also be applied as a comprehensive and practical predictive model for the other rare genetic diseases.

1. Introduction

Rare diseases are a series of diseases that are often chronic and result in lifelong disability or early death [1]. It was estimated that about 400 million people all over the world suffered from rare diseases [2,3], including 16.8 million patients in China [4]. Children

* Corresponding author.
** Corresponding author.
E-mail addresses: lichen_bk@fudan.edu.cn (L. Chen), sunguiqin2001@163.com (G. Sun).
[1] These authors contributed equally: Shuying Yuan, Yanwen Chen.

https://doi.org/10.1016/j.heliyon.2024.e28787
Received 1 August 2023; Received in revised form 23 March 2024; Accepted 25 March 2024
Available online 6 April 2024

were the most affected by rare diseases which were severe multisystem disorders [5], an important cause of severe pediatric mortality [3]. On account of the variable phenotypes, it was difficult for clinicians to diagnose rare disease patients [4,6]. Next-generation sequencing technologies, especially whole-exome sequencing and whole-genome sequencing, could help clinicians increase the discovery rate of causative genes and improve the diagnosis of rare diseases [7,8]. However, there were differences between genotypes and phenotypes, and most rare diseases still lack specific treatments [9]. Patients with rare diseases need more attention from society, and research on rare diseases needs to be continuously promoted by researchers. Research on the accurate and rapid diagnostic assay, effective clinical therapy, and scientific prognosis of rare diseases needed to be carried out.

The human cytosolic peptide:N-glycanase (PNGase; N-glycanase 1, NGLY1), encoded by *NGLY1* gene (OMIM* 610661), acts on N-glycoproteins to generate free N-glycans and deglycosylated peptides [10]. *NGLY1* mutation causes a congenital disorder of deglycosylation (NGLY1-CDDG), a rare human autosomal recessive disease [11,12]. More than 100 NGLY1-CDDG patients were found worldwide, including 6 cases reported in China [13–15] (NGLY1 deficiency handbook, www.NGLY1.org). NGLY1-CDDG is a multisystemic neurodevelopmental disorder in which the individuals exhibit typical features, such as developmental delay, multifocal epilepsy, microcephaly, intellectual disability, liver dysfunction, alacrima, hypohidrosis, and movement disorder [14,16]. NGLY1-CDDG has attracted more attention from scholars and organizations, especially NGLY1.org which was a leading patient support and research organization for NGLY1 Deficiency (www.NGLY1.org).

However, identifying NGLY1-CDDG is a challenge for clinicians owing to the lack of a specific clinical phenotype [13]. Previously, NGLY1-CDDG patients were diagnosed based on gene sequencing data [11,14,17]. Two potential biomarkers of NGLY1-CDDG have been described, including the increased aspartyl-glycosamine content in blood reported by Haijes et al. [13] and the N-glycoprotein structure of Neu5Ac1Hex1GlcNAc1-Asn detected in urine reported by Hall et al. [18]. In 2022, an assay method was used for measuring the enzymatic activity of recombinant and endogenous NGLY1 [15].

NGLY1-CDDG patients harbor various *NGLY1* mutations. Mutations in the *NGLY1* gene might affect the biological function of NGLY1 in vivo, leading to disease. Our group reported, for the first time, a mutation in NGLY1 wherein arginine-328 (R328) was mutated to glycine (R328G), which was a novel functional, active site in NGLY1 [14]. We also found an arginine to leucine mutation at 328 site (R328L) in another NGLY1-CDDG patient (unpublished). Recently, Dabaj et al. reported an NGLY1-CDDG patient harboring an arginine to cysteine mutation (R328C) [19]. We analyzed the amino acid sequence of NGLY1 from *Homo sapiens*, *Pan troglodytes*, *Mus musculus*, *Danio rerio, Gallus gallus, Drosophila melanogaster, Caenorhabditis elegans*, *Saccharomyces cerevisiae*, and *Schizosaccharomyces pombe* and showed that the NGLY1-R328 site was conserved. Furthermore, Katiyar et al. reported that the R210 site (corresponding to position R328 of NGLY1) of the yeast peptide:*N-glycanase* (Png1p) was important for the folding and structural stability of the enzyme [20]. Therefore, we speculated that the NGLY1-R328 site mutating to other amino acids might be a potential pathogenic mutation leading to NGLY1-CDDG. However, it was unclear whether these mutations would cause the dysfunction in NGLY1; thus, a relationship between genotype (*NGLY1* mutation) and phenotype (NGLY1 activity) needs to be established.

In this study, a novel effective model of NGLY1-glycopeptide interaction was simulated to predict the potential pathogenic mutations, and an in vitro standard assay was established for determining the enzymatic activity of mutations reported in NGLY1-CDDG patients. Considering R328 as an example, full-spectrum amino acid mutation libraries of the R328 site to the other 19 amino acids were constructed. The novel model between NGLY1 and glycopeptide and the s in vitro assay might provide a quick screening and diagnosis for all mutations in NGLY1, and a basis for basic clinical research of NGLY1 drug development and molecular pathogenesis.

2. Material and methods

2.1. Bioinformatics analysis of NGLY1

Amino acid sequence blast of NGLY1 and other eukaryotic PNGase. Protein sequences for NGLY1 and other ten eukaryotic PNGase, including *Pan troglodytes*, *Mus musculus*, *Gallus gallus*, *Danio rerio*, *Oryza sativa Japonica Group*, *Arabidopsis thaliana*, *Drosophila melanogaster*, *Caenorhabditis elegans*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, used in the multiple-sequence alignment were downloaded from GenBank™. The.pdb files of proteins with known structures were obtained from the Protein Data Bank. The protein structure prediction of NGLY1 and its mutations was performed using AlphaFold2 [21,22]. In addition, the amino acid side chains of NGLY1 and its mutations were analyzed using the PyMOL software. Multiple sequence alignments were performed using CLUSTALW, and the figure was produced using ESPript 3.0 [23].

2.2. A novel model of the interaction between NGLY1 and glycopeptide

Based on the sequence of amino acids, the protein structure prediction of NGLY1 and its mutations was performed using AlphaFold2 (https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb?pli=1#scrollTo=kOblAoxetgx) [21,22]. Hydrogen atoms were added, and the missing amino acids were edited through the Protein preparation wizard module in Schrodinger [24].

The glycopeptides were acted as docking ligands in this study. The structure of (GlcNAc)2(Man)3(Asn)1 (CID: 52921603) was referenced at PubChem. Based on the structure of (GlcNAc)2(Man)3(Asn)1, the other structures were drawn by ChemDraw, including (GlcNAc)2(Man)2(Asn)1, (GlcNAc)2(Man)1(Asn)1, (GlcNAc)2(Asn)1, and (GlcNAc)1(Asn)1. Those structures were submitted for protonation and energy minimization calculations using the LigPrep module in Schrodinger [24].

NGLY1 comprises three domains: PUB, PNG Core, and PAW, in which the PNG Core domain is the core catalytic domain [25]. A receptor pocket was generated in the Receptor Grid Generation module with the PNG Core domain as the center (the central location

was x = 5.47, y = 4.28, z = 6.22) as the collection of amino acids enclosed within an 20 Å radius sphere centered on the bound ligand. The compounds were screened for NGLY1 binding properties using semi-flexible ligands in the internal coordinate's space. The NGLY1/Glycopeptide complex models were considered for virtual screening using Glide Standard-precision (SP) and Extra-precision (XP) [26].

2.3. Established an in vitro standard assay of NGLY1 enzymatic activity

To determine NGLY1 activity in vitro, the NGLY1 protein was expressed and purified using a standard method [14,27]. The protocol of recombinant protein purification was referenced as previously reported [14]. *NGLY1* cDNA was cloned into the pET-28a (+) plasmid. The recombinant plasmid was verified by sequencing and transformed into *Escherichia coli* BL21 (DE3) cells (Tiangen, China). The transformants were cultured at 37 °C for 12–16 h in Luria-Bertani (LB) medium (Binhe, China) containing 50 μg/mL kanamycin (Sangon Biotech, China) to obtain isolated colonies. IPTG (1.0 mM) was used to induce the expression of the protein at 28 °C for 12 h. The cells were harvested by centrifugation at 3500 rpm 10 min and the pellet was resuspended in a lysis buffer (10 mM imidazole, 20 mM Tris-HCl, 500 mM NaCl, pH 7.4). The suspended cells were sonicated (program: 5 s of sonication followed by a pause (10 s) per cycle, 120 W of power) for 30 min on ice. The debris was removed by centrifugation (30 min, 12,000 rpm) at 4 °C. The supernatant was loaded on a HisPur™ Ni-NTA Spin Columns (Thermo Scientific). The expressed protein on loaded column was washed three times with a wash buffer (25 mM imidazole, 20 mM Tris, 500 mM NaCl, pH 7.4), and the eluted in an elution buffer (250 mM imidazole, 20 mM Tris, 500 mM NaCl, pH 7.4). Mutations of NGLY1 and PNGase F were expressed and purified respectively with the same method. The concentration of NGLY1 and its mutations, and PNGase F was determined by NanoDrop™ One/One C Microvolume UV–Vis Spectrophotometer (Thermo Scientific).

The standard N-glycoprotein bovine pancreatic ribonuclease B (RNase B, New England Biolabs, USA) was used as a substrate to measure NGLY1 activity. Oligosaccharides of RNase B cleaved by NGLY1 were analyzed using MALDI-TOF mass spectrometry (MS) and capillary electrophoresis (CE) [27–30]. For MALDI-TOF MS analysis, glycan profiling was performed in the positive ion reflectron mode in a rapifleX MALDI-TOF mass spectrometer (Bruker, Germany), using a nitrogen-pulsed laser (337 nm) and an acceleration voltage of 20 kV [27,28]. For CE analysis, the sample was derivatized with 8-amino-1,3,6-pyrenetrisulfonic acid (APTS; Molecular Probes, USA). The labeled N-glycans were analyzed through DNA sequencer-assisted, fluorophore-assisted carbohydrate electrophoresis (DSA-FACE) technology using a CE-based ABI 3500 Genetic Analyzer (Applied Biosystems, USA). Data was analyzed using the GeneMapper software version 4.1 [29].

The N-glycan profiles of serum from healthy controls were performed following the instructions of the Glycan-Test Kit (Sysdiagno Biomedtech, China), as previously reported [30,31].

The enzyme digestion reaction was performed in a 20 μL mixture containing 2 μg of RNase B, 2 μg of NGLY1, and PBS added up to 20 μL at 37 °C in a water bath for 12 h. This protocol describes preparation and conducting of SDS-PAGE gels, followed by staining to detect proteins using Coomassie Brilliant Blue (CBB). Then the stained SDS-PAGE gels were scanned to images for subsequent analysis. The following factors were evaluated in the abovementioned reaction conditions: (1) Substrate specificity - the following specific substrates were used: RNase B with N-linked high-mannose oligosaccharide, ovalbumin (OVA, LABLEAD, China) with N-linked hybrid oligosaccharide (LABELED, China), and self-made immunoglobulin G (IgG) with N-linked complex oligosaccharide. The protocol of self-made IgG was referenced as previously reported [65]. The gene was subcloned to pOptiVEC-TOPO and pcDNA 3.3-TOPO (Freedom DG44 Kit, Invitrogen) and expressed in CHO/DHFR G44 cells, IgG was purified through a G protein chromatography and polishing protocol.

(2) Effects of temperature - NGLY1 and RNase B were mixed and incubated at the temperature from 4 °C to 85 °C to determine the effect of the reaction temperature;
(3) Thermal stability - NGLY1 was preincubated at 4 °C, 25 °C, 35 °C, 45 °C, 55 °C, 65 °C, 75 °C, and 85 °C for 1 h prior to the addition of RNase B.
(4) Effects of pH - NGLY1was incubated at varying pH values from 3.5 to 9.0.
(5) Effect of chemical compounds - The reaction mixture was supplemented with additional chemical compounds in a 37 °C water bath for 12 h. N-acetyl-L-cysteine (NAC: 0.1, 0.5, 1.0, and 5 mM), dithiothreitol (DTT: 5, 10, 20, 30, 40, and 50 mM), sodium dodecyl sulfate (SDS, 0.1 % and 0.5 %), benzyloxycarbonyl-Val-Ala-Asp-fluoromethyl ketone (Z-VAD-FMK, 0.1 and 0.5 mM) and hypochlorous acid (0.001 % and 0.005 %).

2.4. Vector construction of NGLY1-CDDG mutations and full-spectrum amino acid mutation libraries on R328 site mutations of PNG core domain of NGLY1

The protocol of construction vectors of nonsense mutations and frameshift mutations were referenced by the method of recombinant plasmid. Primers for mutations were designed using SnapGene application (Table S1). The mutations were obtained using a PCR-based mutagenesis protocol. The PCR product was ligated between the *Nco*I (Thermo Fisher Scientific, USA) and *Xho*I (Thermo Fisher Scientific, USA) cloning sites on the pET28a plasmid.

The protocol of construction vectors of missense mutations and full-spectrum amino acid mutation libraries on R328 site was obtained using a PCR-based mutagenesis protocol following previously reported studies [14]. Primers for mutations were designed using QuikChange Primer Design (https://www.agilent.com/store/primer Design Program.jsp) (Table S1 and Table S2).

All the mutated plasmids were confirmed by sequencing and transformed into *E. coli* BL21 (DE3) cells for protein expression.

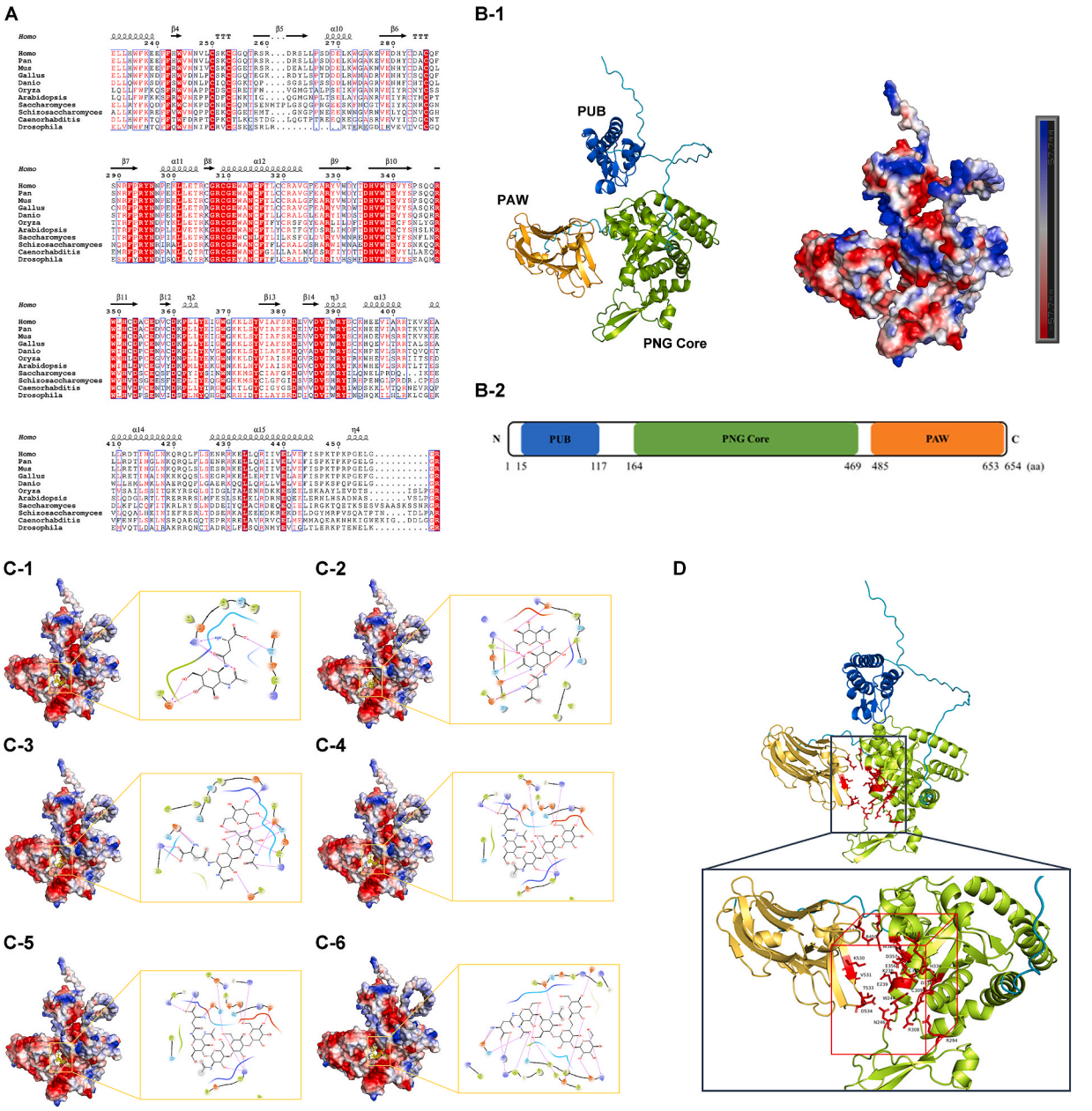
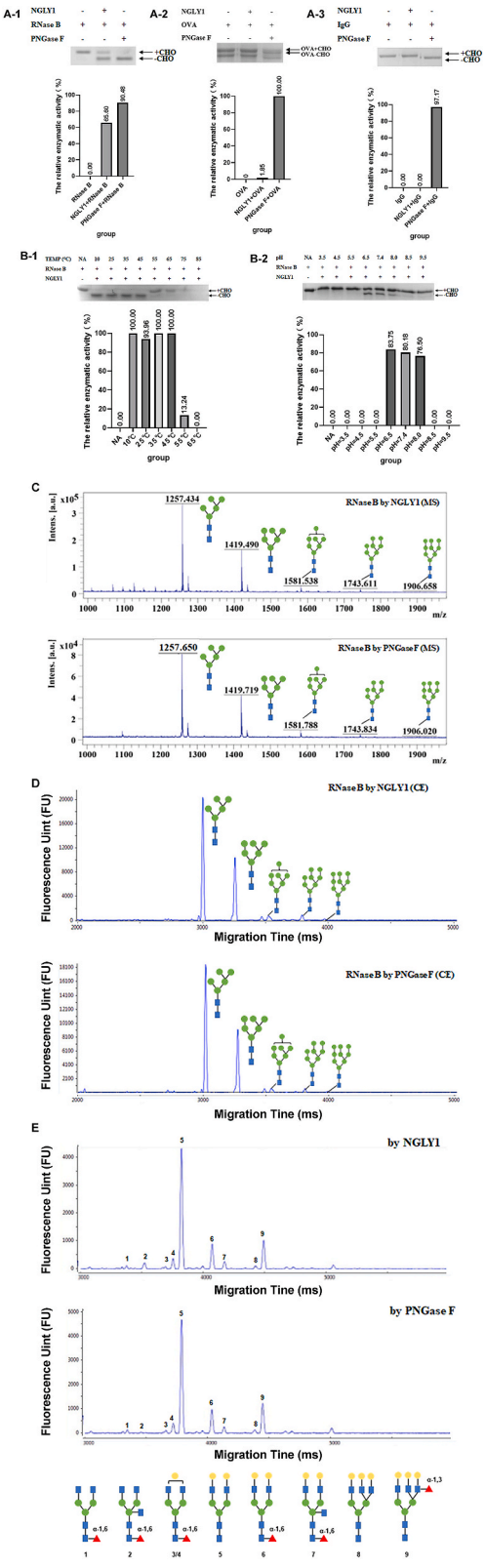


Fig. 1. The simulation structure prediction and predicted binding sites between NGLY1 and glycopeptide. (A) Sequence analysis of NGLY1. Structure-based sequence alignment of some eukaryotic PNGase. The amino acid sequences of PNGase of *Homo sapiens* (Homo), *Pan troglodytes* (Pan), *Mus musculus* (Mus), *Gallus gallus* (Gallus), *Danio rerio* (Danio), *Oryza sativa Japonica Group* (Oryza), *Arabidopsis thaliana* (Arabidopsis), *Saccharomyces cerevisiae* (Saccharomyces), *Schizosaccharomyces pombe* (Schizosaccharomyces), *Caenorhabditis elegans* (Caenorhabditis), and *Drosophila melanogaster* (Drosophila) were aligned by CLUSTALW. The 3D structures of NGLY1 were predicted by AlphaFold2, and the figure was produced using ESPript. (B-1) The simulated structure of NGLY1 by AlphaFold2. (B-2) Electrostatic surface potential diagrams of NGLY1. The color red was negative, while the color blue was positive. (C-1) The three-dimensional structure of predicted binding sites between NGLY1 and (Asn)-GlcNAc. (C-2) The three-dimensional structure of predicted binding sites between NGLY1 and (Asn)-GlcNAc-(1–4)-GlcNAc. (C-3) The three-dimensional structure of predicted binding sites between NGLY1 and (Asn)-Man-(1–4)-GlcNAc-(1–4)-GlcNAc. (C-4) The three-dimensional structure of predicted binding sites between NGLY1 and (Asn)-Man-(1–3)-Man-(1–4)-GlcNAc-(1–4)-GlcNAc. (C-5) The three-dimensional structure of predicted binding sites between NGLY1 and (Asn)-Man-(1–6)-Man-(1–4)-GlcNAc-(1–4)-GlcNAc. (C-6) The three-dimensional structure of predicted binding sites between NGLY1 and (Asn)-Man-(1–3)-[Man-(1–6)]-Man-(1–4)-GlcNAc-(1–4)-GlcNAc. (D) The binding site between NGLY1 and glycopeptide. The red box presented the simulated pocket of potential pathogenic sites.



(caption on next page)

Fig. 2. The enzymatic activity of recombinant NGLY1. Each experiment was done once (n = 1). (A-1) SDS-PAGE analysis of RNase B treated with NGLY1. (A-2) SDS-PAGE analysis of OVA treated with NGLY1. (A-3) SDS-PAGE analysis of IgG treated with NGLY1. The schematic diagram of RNase B, OVA, and IgG is shown [27,52,53]. RNase B: ribonuclease B; OVA: ovalbumin; IgG: immunoglobulin G; CHO: carbon hydrogen oligosaccharide; +CHO: glycosylated forms of the substrates; –CHO: deglycosylated forms of the substrates. (B-1) The enzymatic reaction temperature of NGLY1. (B-2) Enzymatic reaction pH of NGLY1. NA: substrate control, without treatment; "-": without NGLY1. (C) The released N-glycans from RNase B were detected by MALDI-TOF MS. (D) The released N-glycans from RNase B were detected by CE [36]. (E) The released N-glycans from total serum glycoproteins of healthy individuals were detected by CE. Structure formulas: blue square, N-acetylglucosamine; green circle, mannose; yellow circle, galactose; red triangle, fucose.

2.5. Enzymatic activity assay of NGLY1 mutations

RNase B was chosen as a substrate for measuring the activity of mutated NGLY1 and analyzed by SDS-PAGE and ImageJ software. The reaction condition for qualitative enzymatic activity of NGLY1 mutations was referred to the wild-type NGLY1. For the relative enzymatic activity, 1 μg RNase B and 2 μg of mutated NGLY1 were mixed in a 20 μL reaction system. The reaction was conducted at 37 °C in a water bath for 1 h or 12 h.

3. Results

3.1. Building a novel model of the interaction between NGLY1 and glycopeptides

A structure-based sequence alignment was conducted to predict the active sites of NGLY1 (Fig. 1). There were 38 conserved amino acid residues between 240 aa and 400 aa, Fig. 1A). Among them, there were 8 negatively charged amino acids and 6 positively charged amino acids. It could be speculated that the electrification of amino acids would affect the structure and function of NGLY1.

NGLY1 comprises three domains: PUB, PNG Core, and PAW. The PUB domain is associated with the ubiquitin-proteasome system, PAW is involved in the oligosaccharide binding of PNGase, and the PNG Core domain is the core catalytic domain [25]. The catalytic domain of NGLY1 (PNG Core) is also known as Transglutaminase domain [15,32,33]. Based on the simulated structure of NGLY1 by AlphaFold2, the ranges of amino acid for the three domains are PUB domain (15–116 aa), PNG Core domain (164–469 aa), and PAW domain (485–653 aa), respectively (Fig. 1B).

N-glycosylation is the attachment of N-glycans to a glycosylation sequin (N-X-S/T, where X cannot be Proline), and the unique branch structure of each N-glycan was a (GlcNAc)2(Man)3 core [34]. The substrate-binding region was identified by the superimposing NGLY1 structure with a crystal structure of glycan product, with a binding domain formed by PNG Core and PAW (Fig. 1C). Two binding sites were speculated: the binding sites for glycans which were constituted with Glu239, Asn246, Cys309, Cys355, Asp357, Arg458, Lys530, Val531, Asp534, Glu544; the conjunction sites of peptide and glycans which was constituted with Lys238, Trp244, Arg294, Arg308, Asp335, His336, Glu356, Trp369, Lys371, Thr533. The simulated pocket of potential pathogenic sites consisted of 66 amino acids, including amino acids 235 to 246, 277 to 279, 292 to 295, 305 to 316, 332 to 338, 351 to 359, 365 to 374, and 452 to 460. The predicted range of missense mutation was reduced from the domain of PNG Core contained 306 amino acids to the simulated pocket contained 66 amino acids by this model (Fig. 1D). It was worth noting that most amino acids were charged, which might suggest that the charge of amino acids were functionally related.

3.2. Mutated sites of NGLY1-CDDG patients and their enzymatic activity

To analyze the enzymatic activity of mutated NGLY1, an in vitro enzymatic assay needed to be established. The results revealed that NGLY1 could prefer to release N-linked glycans from RNase B (Fig. 2A and Fig. S4A). The enzymatic activity was maintained at the reaction temperature range of 10 °C–45 °C (Figs. 2B–1 and Fig S4B-1). NGLY1 was active at pH between 6.5 and 8.0 (Fig. 2B- and Figs. S4B–2). Based on these results, the optimum reaction conditions for NGLY1 were pH 7.4, temperature 37 °C, and substrate RNase B.

The profile of the released glycans from RNase B treated with NGLY1 was detected through MS (Fig. 2C), consistent with the reported N-glycan structures [27,35]. The results of CE also showed that the profiles of glycans released from RNase B were the same as those after treatment with NGLY1 or PNGase F (Fig. 2D) [36]. These results indicated that the recombinant NGLY1 could hydrolyze the intact glycans from RNase B.

The function of NGLY1 has been proven by the release of N-glycans on healthy serum glycoproteins through the CE method (Fig. 2E). In addition, the N-glycome profiling from serum glycoproteins was similarly treated with NGLY1 and PNGase F, which was referenced from previous reports [30,31]. NGLY1 could catalyze N-glycopeptides with a core-a-1,6-fucosylated (Fig. 2E).

According to reports, there are more than 100 cases of NGLY1-CDDG patients worldwide [13–15]. A total of 30 exon mutations reported in NGLY1-CDDG patients were collected (Table 1), which were also revealed to not be SNPs by gnomAD database (http://gnomad-sg.org/) and "HUABIAO" whole-exome public database [37]. At present, *NGLY1* mutations in NGLY1-CDDG patients mainly occurred in exon, and the site with the highest mutation frequency was c.1201A > T (p.R401X); the most varied nonsense mutation amino acid site was R328, including R328C, R328G, and R328L. For those nonsense mutations and frameshift mutations, the structure of mutations was changed, which might affect the enzymatic activity. Interestingly, missense mutations occurred in NGLY1-CDDG patients, in which most happened beside the predicted NGLY1-glycopeptide binding region (Fig. 1C and D).

The enzymatic activity of 30 mutated NGLY1 was evaluated using RNase B as a substrate. Wild type NGLY1 was used as a positive

Table 1
Mutations in exons of *NGLY1* in reported NGLY1-CDDG patients.

Patient number	Nucleotide change (NM_018297)	Amino-acid change	Mutated remained structure (remained amino acids)	Relative enzymatic activity (%)[a]	Clinical phenotypes						
					Global developmental delay	Small hand and feet	Microcephaly	Neuromotor developmental delay/ abnormalities	alacrima/ hypohidrosis	Ocular dysfunction	liver dysfunction
P1 [11]	c.1201A > T	p.R401X	PUB + partial PNG Core (400 aa)	0.00	Yes	–	–	Yes	Yes	Yes	Yes
	c.1891delC	p.Q631HfsX7	PUB + PNG Core + partial PAW (636 aa)	11.41							
P2 [12]	c.1201A > T	p.R401X	PUB + partial PNG Core (400 aa)	0.00	Yes	Yes	No	Yes	Yes	Yes	Yes
	c.1891delC	p.Q631HfsX7	PUB + PNG Core + partial PAW (636 aa)	11.41							
P3 [12]	c.1370dupG	p.R458KfsX14	PUB + partial PNG Core (470 aa)	0.00	Yes	No	Yes	Yes	Yes	Yes	Yes
P4 [12]	c.1205_1207delGAA	p.402_403delR	PUB + PNG Core + PAW (653 aa)	41.87	Yes	Yes	Yes	Yes	Yes	Yes	Yes
P5 [12]	c.1201A > T	p.R401X	PUB + partial PNG Core (400 aa)	0.00	Yes	Yes	No	Yes	Yes	Yes	Yes
P6 [12]	c.1201A > T	p.R401X	PUB + partial PNG Core (400 aa)	0.00	Yes	No	Yes	Yes	Yes	Yes	Yes
P7 [12]	c.1201A > T	p.R401X	PUB + partial PNG Core (400 aa)	0.00	Yes	No	Yes	Yes	–	–	Yes
P8 [12]	c.1201A > T	p.R401X	PUB + partial PNG Core (400 aa)	0.00	Yes	No	Yes	Yes	Yes	Yes	Yes
P9 [12]	c.1201A > T	p.R401X	PUB + partial PNG Core (400 aa)	0.00	Yes	No	Yes	Yes	Yes	Yes	–
P10 [54]	c.1533_1536delTCAA	p.N511KfsX51	PUB + PNG Core + partial PAW (560 aa)	0.00	Yes	No	–	Yes	Yes	Yes	Yes
P11 [54]	c.1533_1536delTCAA	p.N511KfsX51	PUB + PNG Core + partial PAW (560 aa)	0.00	Yes	No	–	Yes	Yes	Yes	–
P12 [55]	c.347C > G	p.S116X	PUB (115 aa)	0.00	Yes	No	Yes	Yes	Yes	Yes	Yes
P13 [56]	c.1201A > T	p.R401X	PUB + partial PNG Core (400 aa)	0.00	Yes	–	–	Yes	–	Yes	–

(*continued on next page*)

Table 1 (*continued*)

Patient number	Nucleotide change (NM_018297)	Amino-acid change	Mutated remained structure (remained amino acids)	Relative enzymatic activity (%)[a]	Clinical phenotypes						
					Global developmental delay	Small hand and feet	Microcephaly	Neuromotor developmental delay/ abnormalities	alacrima/ hypohidrosis	Ocular dysfunction	liver dysfunction
P14 [57]	c.953T > C	p.L318P	PUB + PNG Core + PAW (654 aa)	0.00	Yes	–	–	Yes	–	Yes	Yes
	c.1169G > A	p.R390P	PUB + PNG Core + PAW (654 aa)	92.13							
P15 [57]	c.730T > C	p.W244R	PUB + PNG Core + PAW (654 aa)	0.00	Yes	–	–	Yes	–	–	–
	c.931G > A	p.E311K	PUB + PNG Core + PAW (654 aa)	12.24							
P16 [57]	c.1604G > A	p.W535X	PUB + PNG Core + partial PAW (534 aa)	0.00	Yes	–	–	Yes	Yes	Yes	–
	c.1910delT	p.L637X	PUB + PNG Core + partial PAW (636 aa)	14.98							
P17 [57]	c.622C > T	p.Q208X	PUB + partial PNG Core (207 aa)	0.00	Yes	–	–	Yes	–	Yes	–
P18 [57]	c.622C > T	p.Q208X	PUB + partial PNG Core (207 aa)	0.00	Yes	–	–	Yes	–	Yes	–
P19 [57]	c.1201A > T	p.R401X	PUB + partial PNG Core (400 aa)	0.00	Yes	–	–	Yes	Yes	Yes	–
P20 [13]	c.1756C > T	p.R586X	PUB + PNG Core + partial PAW (585 aa)	19.44	Yes	–	Yes	Yes	Yes	Yes	Yes
P21 [13]	c.1756C > T	p.R586X	PUB + PNG Core + partial PAW (585 aa)	19.44	Yes	–	Yes	Yes	Yes	Yes	Yes
P22 [13]	c.1853T > G	p.L618X	PUB + PNG Core + partial PAW (617 aa)	24.84	–	–	Yes	Yes	–	Yes	Yes
P23 [16]	c.1201A > T	p.R401X	PUB + partial PNG Core (400 aa)	0.00	Yes	–	–	Yes	Yes	–	Yes
P24 [58]	c.1405C > T	p.R469X	PUB + PNG Core (468 aa)	64.23	Yes	–	Yes	Yes	Yes	–	Yes
P25 [59]	c.1837delA	p.I613FfsX6	PUB + PNG Core + partial PAW (617 aa)	0.00	Yes	–	–	Yes	–	–	–

(*continued on next page*)

Table 1 (*continued*)

Patient number	Nucleotide change (NM_018297)	Amino-acid change	Mutated remained structure (remained amino acids)	Relative enzymatic activity (%)[a]	Clinical phenotypes						
					Global developmental delay	Small hand and feet	Microcephaly	Neuromotor developmental delay/ abnormalities	alacrima/ hypohidrosis	Ocular dysfunction	liver dysfunction
P26 [17]	c.531dupC	p.N178QfsX9	PUB + partial PNG Core (185 aa)	0.00	Yes	–	–	Yes	Yes	–	Yes
	c.1891delC	p.Q631HfsX7	PUB + PNG Core + partial PAW (636 aa)	11.41							
P27 [60]	c.1063T > C	p.C355R	PUB + PNG Core + PAW (654 aa)	0.00	Yes	–	–	Yes	Yes	–	Yes
P28 [61]	c.849T > G	p.C283W	PUB + PNG Core + PAW (654 aa)	0.00	Yes	Yes	–	Yes	–	Yes	Yes
	c.1201A > T	p.R401X	PUB + partial PNG Core (400 aa)	0.00							
P29 [61]	c.1201A > T	p.R401X	PUB + partial PNG Core (400 aa)	0.00	Yes	Yes	–	Yes	Yes	Yes	Yes
P30 [61]	c.1201A > T	p.R401X	PUB + partial PNG Core (400 aa)	0.00	Yes	Yes	–	Yes	–	Yes	–
	c.1067A > G	p.E356G	PUB + PNG Core + PAW (654 aa)	0.00							
P31 [61]	c.1837delA	p.I613FfsX6	PUB + PNG Core + partial PAW (617 aa)	0.00	Yes	No	–	Yes	–	–	–
P32 [62]	c.1891delC	p.Q631HfsX7	PUB + PNG Core + partial PAW (636 aa)	11.41	Yes	Yes	No	Yes	–	–	Yes
P33 [63]	c.1156G > T	p.D386Y	PUB + PNG Core + PAW (654 aa)	0.00	Yes	–	–	Yes	–	–	Yes
	c.1168C > T	p.R390X	PUB + partial PNG Core (389 aa)	0.00							
P34 [14]	c.1025A > G	p.Y342C	PUB + PNG Core + PAW (654 aa)	82.49	Yes	No	Yes	Yes	Yes	–	Yes
	c.1231C > T	p.R411X	PUB + partial PNG Core (410 aa)	0.00							
P35 [14]	c.1025A > G	p.Y342C	PUB + PNG Core + PAW (654 aa)	82.49	Yes	No	Yes	Yes	Yes	–	Yes

(*continued on next page*)

Table 1 (*continued*)

Patient number	Nucleotide change (NM_018297)	Amino-acid change	Mutated remained structure (remained amino acids)	Relative enzymatic activity (%)[a]	Clinical phenotypes						
					Global developmental delay	Small hand and feet	Microcephaly	Neuromotor developmental delay/ abnormalities	alacrima/ hypohidrosis	Ocular dysfunction	liver dysfunction
	c.1025A > G	p.Y342C	PUB + PNG Core + PAW (654 aa)	82.49							
P36 [14]	c.1025A > G	p.Y342C	PUB + PNG Core + PAW (654 aa)	82.49	Yes	No	Yes	Yes	Yes	–	Yes
	c.1025A > G	p.Y342C	PUB + PNG Core + PAW (654 aa)	82.49							
P37 [14]	c.1637-1652del CATCTTTTGCTTATAT	p.S546FfsX12	PUB + PNG Core + partial PAW (556 aa)	15.61	Yes	Yes	Yes	Yes	Yes	Yes	Yes
P38 [14]	c.1637-1652del CATCTTTTGCTTATAT	p.S546FfsX12	PUB + PNG Core + partial PAW (556 aa)	15.61	Yes	Yes	Yes	Yes	Yes	Yes	Yes
P39 [14]	c.982C > G	p.R328G	PUB + PNG Core + PAW (654 aa)	0.00	Yes	Yes	Yes	Yes	Yes	Yes	Yes
P40 [19]	c.982C > T	p.R328C	PUB + PNG Core + PAW (654 aa)	14.95	Yes	–	Yes	Yes	–	Yes	Yes
P41 [64]	c.708G > T	p.W236C	PUB + PNG Core + PAW (654 aa)	0.00	Yes	No	No	Yes	–	–	No
P42 [64]	c.708G > T	p.W236C	PUB + PNG Core + PAW (654 aa)	0.00	Yes	No	No	Yes	–	–	No
P43 [64]	c.708G > T	p.W236C	PUB + PNG Core + PAW (654 aa)	0.00	Yes	No	No	Yes	–	–	–

[a] The reaction time: 12 h; - not available or not applicable.

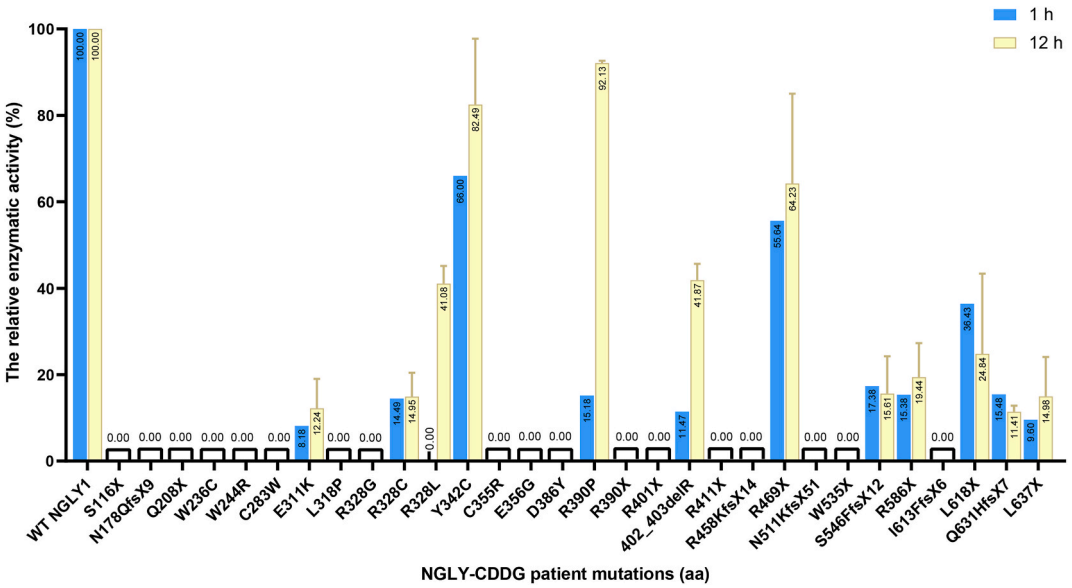


Fig. 3. The relative enzymatic activity of mutations reported in NGLY1-CDDG patients. The error bars depicted the results of three repetitions in the 12-huor reaction time (3 batches of NGLY1 and its mutations). The data are presented as mean ± SEM.

control. It was shown that 18 mutations were dysfunctional (Fig. 3 and Fig. S5). For the 12 remained activity mutations, the efficiency of enzymatic activity was lower than that of wild-type NGLY1 (Fig. 3 and Fig. S6). It was speculated that the predicted NGLY1-glycopeptide binding sites were important to the function of NGLY1. However, the established model was not perfect. Several mutated sites occurred outside the NGLY1-glycopeptide binding region, including C283, E311, L318, R328, Y342, D386, and R390, which might suggest that the charged amino acids near the binding region also influenced the enzymatic activity.

3.3. Enzymatic activities of NGLY1 functional active site R328 mutations

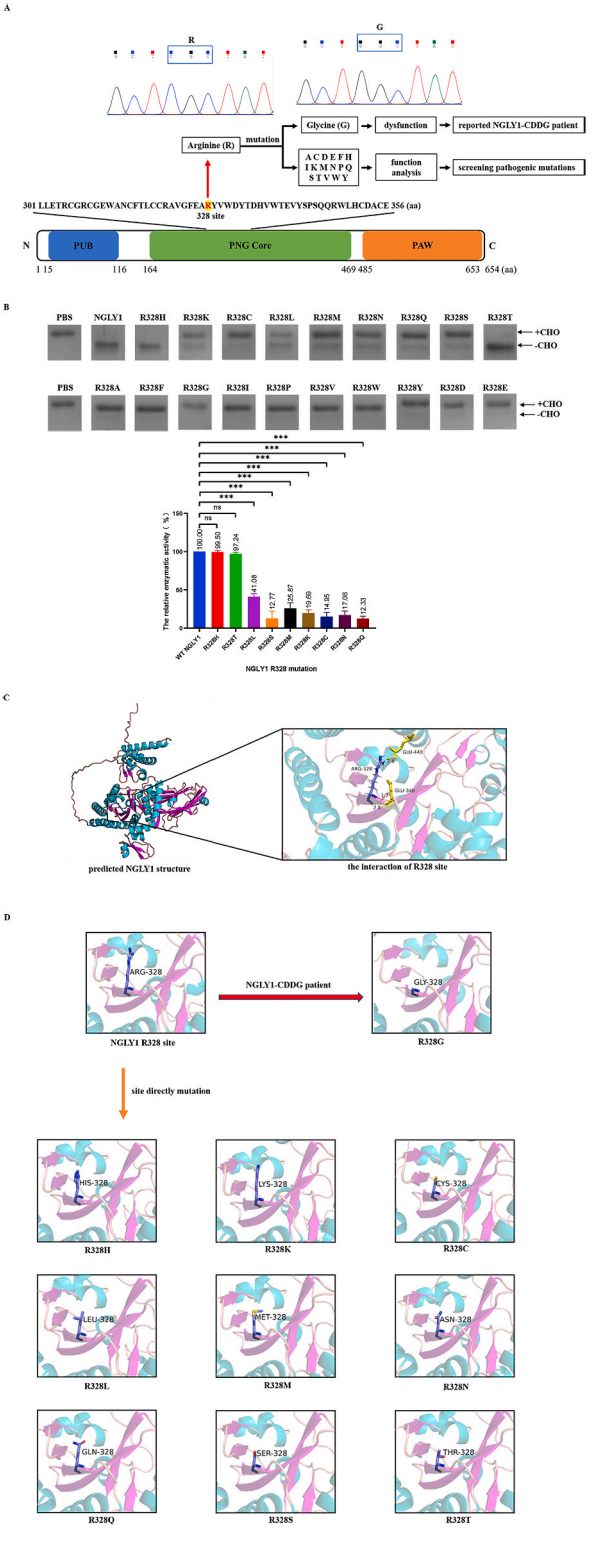
According to the analysis of NGLY1-CDDG patient mutations, the disease also happened when the same amnio acid site mutated to the others, such as the R328 site of NGLY1 which was a conserved site (Fig. 1A). Considering the R328 site as an example, a protocol of establishing full-spectrum amino acid mutation libraries was constructed. A schematic diagram of R328 NGLY1 mutations was presented in Fig. 4A. The results showed that 10 mutations were dysfunctional, including R328A, R328D, R328E, R328F, R328G, R328I, R328P, R328V, R328W, and R328Y, whereas NGLY1 mutants with R328H and R328T remained active as wild-type NGLY1. The other 7 mutations showed decreasing activity (Fig. 4B).

We also analyzed the 3D structure of mutant NGLY1. The simulated structures of wild-type NGLY1 and mutated NGLY1 (possessing enzymatic activity) predicted by AlphaFold2 were presented in Fig. 4C and D respectively, and the structures of dysfunctional mutations were also predicted (data not shown). Based on the predicted structure, the R328 site was critical for packing the catalytic domain. The R328 site interacted with E340 and E440 (upper part of Fig. 4C), which were conserved residues in NGLY1 and other eukaryotic PNGase (Fig. 1). Disruption of such packing, either at sterically very small (R328A, R328G), large residues (R328F, R328W), or electrification (negatively charged, R328D, R328E) would affect the catalysis. It was speculated that the structure and charge of the amino acid side chain could influence the function of NGLY1.

Based on the full-spectrum amino acid mutation libraries of the NGLY1 R328 site, the charge of amino acids has a great influence on its function, and the charged amino acids around the simulated pocket could also be potential pathogenic sites. It was a modification of the prediction model between NGLY1 and glycopeptides. In the meantime, the method of the full-spectrum mutation library construction could also carry out a more detailed prediction of the reported pathogenic sites.

3.4. Effect of chemical compounds on NGLY1 activity

Since there is no effective treatment for NGLY1-CDDG, the compounds for promoting NGLY1 activity might be detected based on the established in vitro enzymatic assay. In addition, the effects of compounds to the mutations reported by NGLY1-CDDG were also examined. Among them, the results of two mutants were shown in this section, including 402delR [12] and R328L mutations found in an NGLY1-CDDG patient. Several chemical compounds were added to the reaction system, including a reducing agent (NAC and DTT), an enzyme inhibitor (Z-VAD-FMK), and a denaturant (SDS). The results were analyzed by SDS-PAGE (Fig. 5 and Fig. S7). At concentrations of 0.1, 0.5, and 1 mM, NAC promoted the NGLY1 enzymatic activity, while NAC had no significant effect on NGLY1 with 402delRand R328L (Fig. 5A and Fig. S7A). However, enzymatic activity of NGLY1, 402delR- and R328L-carrying NGLY1 was inhibited when NAC was added at a concentration of 5 mM (Fig. 5A and Fig. S7A). NGLY1 activity was increased by DTT, while the activity of the



(caption on next page)

Fig. 4. Function and simulated structure of NGLY1-R328 mutations. (A) The construction of the R328 site in NGLY1. The three domains of NGLY1 were simulated by AlphaFold2. (B) The enzymatic activity of mutations on the R328 site of NGLY1. The error bars depicted the results of three repetitions (3 batches of NGLY1 and its mutations) in the 12-huor reaction time. CHO: carbon hydrogen oligosaccharide; +CHO: glycosylated forms of the substrates; –CHO: deglycosylated forms of the substrates. (C) The interaction of the R328 site with other amino acids (E340 and E440). (D) The predicted structures of mutated NGLY1 (with enzymatic activity). A: alanine; C: cysteine; D: aspartic acid; E: glutamic acid; F: phenylalanine; G: glycine; H: histidine; I: isoleucine; K: lysine; L: leucine; M: methionine; N: asparagine; P: proline; Q: glutamine; R: arginine; S: serine; T: threonine; V: valine; W: tryptophan; Y: tyrosine. The data are presented as mean ± SEM. ***$p < 0.001$, ns: no significant. One-way ANOVA multiple comparisons. The relatively enzymic activity of other 10 dysfunctional mutations were not shown ($P < 0.001$).

NGLY1 mutant harboring 402delR and R328L (Fig. 5B and Fig. S7B) was inhibited. The enzymatic activity of NGLY1 was inhibited by Z-VAD-FMK and SDS (Fig. 5C and Fig. S7C). In general, Z-VAD-FMK inhibits pan-caspase and yeast PNGase [38].

4. Discussion

NGLY1-CDDG, the first reported congenital disorder of deglycosylation, is an extremely rare autosomal recessive disorder caused by pathogenic mutations in the *NGLY1* gene [17]. NGLY1-CDDG patients are routinely diagnosed using gene sequencing [11,13,16,17]. However, the *NGLY1* mutations detected by gene sequencing were not all related to NGLY1-CDDG. Therefore, a standard assay for the correlation between gene mutations and NGLY1 enzymatic activity is necessary to screen and diagnose NGLY1-CDDG.

In this study, a novel effective model of NGLY1-glycopeptide interaction was simulated to preliminarily analyzed the enzymatic mechanism of NGLY1 and predict the potential pathogenic mutations. Then, an in vitro enzymatic assay system was established to evaluate the enzymatic activity of NGLY1 using MALDI-TOF MS, CE, and SDS-PAGE. NGLY1 can cleave intact N-glycans from N-glycoproteins [27]. It also proved that NGLY1 could act on a single core-a-1,6-fucosylated biantennary glycan.

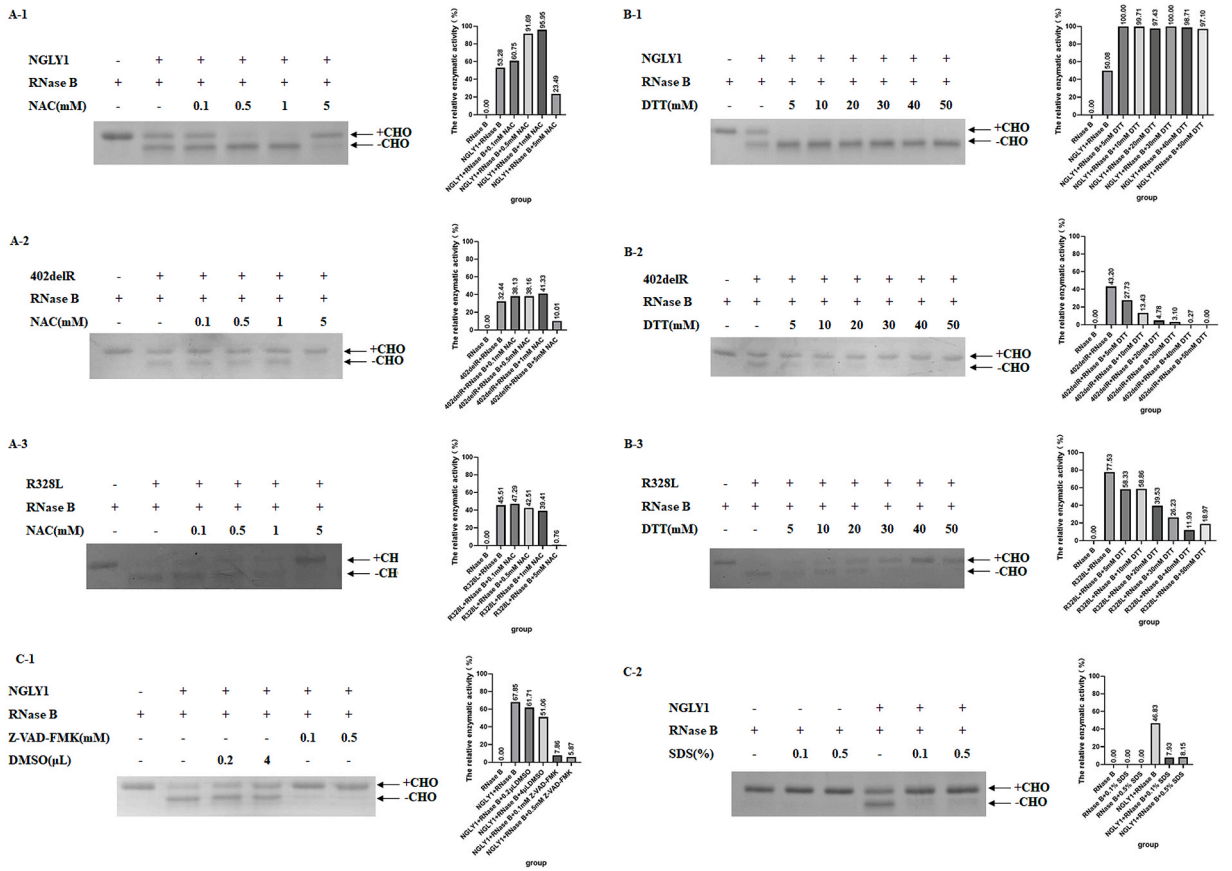


Fig. 5. The enzymatic activity of NGLY1 treated with different chemical compounds. Each experiment was done once (n = 1). (A-1) NGLY1 treated with NAC. (A-2) NGLY1 mutant with 402delR (NGLY1-CDDG mutation) treated with NAC. (A-3) NGLY1 mutant with R328L (NGLY1-CDDG mutation) treated with NAC. (B-1) NGLY1 treated with DTT. (B-2) NGLY1 mutant with 402delR (NGLY1-CDDG mutation) treated with DTT. (B-3) NGLY1 mutant with R328L (NGLY1-CDDG mutation) treated with DTT. (C-1) NGLY1 treated with Z-VAD-FMK. (C-2) NGLY1 treated with SDS. CHO: carbon hydrogen oligosaccharide; +CHO: glycosylated forms of the substrates; –CHO: deglycosylated forms of the substrates.

Combined with the predicted results, five missense mutations happened in the binding region of NGLY1 and glycopeptides. It was proved that the established model of NGLY1 and glycopeptides could be effectively used to explain the catalytic mechanism of NGLY1 and predict the potential pathogenic sites for the screening and diagnosis of NGLY1-CDDG.

There are prokaryotic and eukaryotic cell-free protein synthesis systems. The prokaryotic system has the advantages of high protein yield, simple cultivation and fast cell growth and lysate preparation, cost-efficient, easy genetic engineering, and well-established, but it also has the disadvantage of limited post-translational modifications [39,40]. The eukaryotic system has the advantages of the ability to perform post-translational modifications like glycosylation and correct folding of many protein types [41,42]. Compared with prokaryotic system, eukaryotic system was laborious and expensive lysate preparation, and low protein yield [40]. Therefore, the cell-free protein synthesis in this study was based on *E. coli* cell extracts.

The charged properties of amino acids have a great influence on the activity proved by the full-spectrum mutation libraries of the R328 site mutations on NGLY1. The result illustrated that the positive charge of arginine on 328 is critical for enzymatic activity of NGLY1, as the functional role of R328 could be complemented by electropositive charge, but not the other electronegative amino acids. When R328 was mutated to an amino acid with the same charge (electropositivity), the mutated NGLY1 remained active (including the R328H and R328K), whereas the mutated NGLY1 (electronegative amino acid) was dysfunctional (including the R328D and R328E). For mutations with electroneutrality, the structure of the amino acid side chain influences the function of NGLY1. Besides R328P (328 site at a loop), the 328 site of other 18 mutations were shown at β-sheet. The mutated NGLY1 with straight amino acid side chain showed enzymic activity, including R328C, R328L, R328 M, R328 N, R328Q, R328S, and R328T. The mutated NGLY1 was dysfunctional when the amino acid consisted with benzene ring structure (R328F, R328W, R328Y), or less than one carbon (R328A, R328G). Therefore, it was speculated that the electrification and side chain of amino acids possibly influenced the function of NGLY1. Some studies have indicated that electrification [43] and side-chain structure [44] of amino acids would affect the function of proteins. The constructed full-spectrum functional mutation libraries provided a new diagnostic method for proteins mutated at R328 and other sites in NGLY1-CDDG and provided new insights into other single-gene disorders. Subsequently, we will analyze the correlation between the reported mutations in NGLY1-CDDG patients and NGLY1 enzymatic activity and the effect of R328 mutations on cells.

In this study, the effects of several compounds on NGLY1 enzymatic activity were analyzed using an established in vitro enzymatic assay. This in vitro assay may serve as a practical screening strategy for drugs and compounds with potential therapeutic value for NGLY1-CDDG patients. NAC and DTT have been shown to promote the enzymatic activity of wild-type NGLY1. NAC has been used as a mucolytic agent for treating airway muco-obstructive diseases [45], and it might be a potential treatment for psychiatric disorders, including Alzheimer's dementia, drug-induced neuropathy, and progressive myoclonic epilepsy [46]. In addition, the effect of DTT has been proven to mitigate hepatic and renal injury in bile duct-ligated mice and ameliorate hematopoietic and intestinal injury in irradiated mice [47]. Interestingly, the effects of NAC and DTT were not consistent among NGLY1-CDDG patient mutations. This finding is a cautionary warning of the need for precision medicine, in which treatment varies between patients with different NGLY1 mutations.

The assay established in this study has rooms to improve. The enzymatic activity of NGLY1 could be enhanced by EDTA (data not shown), indicating some inhibitory divalent metal ions existed in the assay. In this study, molecular weight changes between glycosylated and deglycosylated RNase B were used to detect the enzymic activity. The purification of NGLY1 was not high enough by this assay, and there might be a contamination of some proteins whose molecular weight are the same as glycosylated and deglycosylated RNase B. Those might impact the detection of enzymic activity. The quantity and purity of NGLY1 used in the assay could be further improved by using eukaryotic expression system, codon-optimization and other measures. Structural proteins were also not eligible for analysis by this system that relies on enzymatic activity. In addition, the assay time could also be improved for real world clinical needs.

More and more attention has been put on rare diseases. Three levels for researching rare diseases were suggested by our group. The first level was to prove the relationship between gene mutation and disease. PNGase, including NGLY1, was involved in a variety of biological processes in vivo. It has been demonstrated that PNGase was involved in the embryonic development of mice [48] and Drosophila [49], and also affected the egg-laying behavior of *Caenorhabditis elegans* [50]. The first patient caused by mutated *NGLY1* was reported in 2012 [11], while it was recognized as a rare autosomal recessive disease in 2014 [51]. The second level was to provide an assay for screening and diagnosis of rare diseases in clinical treatment. Identifying rare diseases is a challenge for clinicians owing to the lack of a relationship between genotype and phenotype. Our group had established a novel effective model of NGLY1-glycopeptide interaction was simulated to predict the enzymatic mechanism of NGLY1, and an in vitro enzymatic assay system was established to evaluate the enzymatic activity of mutated NGLY1. It provided new sights for screening and diagnosis of NGLY1, and the study of other single gene diseases. The third level was to carry on the catalytic mechanism research for the gene therapy or drug treatment of rare diseases. The effect of compounds could be analyzed by in vitro enzymatic assay systems, which provided a reference for the screening and development of potential therapeutic drugs for rare diseases.

Funding

This work was supported by the Natural Science Foundation of Zhejiang Province [grant number LY20C050002]; Zhejiang Chinese Medical University Science Foundation [grant number 2020ZG25]; and National Natural Science Foundation of China [grant number 31600644].

Data availability statement

The data generated for this study is included in article and supplementary material.

CRediT authorship contribution statement

Shuying Yuan: Writing – original draft, Methodology, Formal analysis, Data curation. **Yanwen Chen:** Writing – original draft, Methodology, Formal analysis, Data curation. **Lin Zou:** Software, Methodology. **Xinrong Lu:** Software, Methodology. **Ruijie Liu:** Software, Data curation. **Shaoxing Zhang:** Software, Data curation. **Yuxin Zhang:** Software, Data curation. **Cuiying Chen:** Software. **Dongqing Cheng:** Data curation. **Li Chen:** Writing – review & editing, Supervision, Conceptualization. **Guiqin Sun:** Writing – review & editing, Supervision, Project administration, Conceptualization.

Declaration of competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e28787.

References

[1] H.I. Roessler, N.V.A.M. Knoers, M.M. van Haelst, G. van Haaften, Drug repurposing for rare diseases, Trends Pharmacol. Sci. 42 (2021) 255–267.
[2] K.M. Boycott, A. Rath, J.X. Chong, T. Hartley, F.S. Alkuraya, G. Baynam, A.J. Brookes, M. Brudno, A. Carracedo, J.T. den Dunnen, S.O.M. Dyke, X. Estivill, J. Goldblatt, C. Gonthier, S.C. Groft, I. Gut, A. Hamosh, P. Hieter, S. Hohn, M.E. Hurles, P. Kaufmann, B.M. Knoppers, J.P. Krischer, M. Macek, G. Matthijs, A. Olry, S. Parker, J. Paschall, A.A. Philippakis, H.L. Rehm, P.N. Robinson, P.C. Sham, R. Stefanov, D. Taruscio, D. Unni, M.R. Vanstone, F. Zhang, H. Brunner, M.J. Bamshad, H. Lochmuller, International cooperation to enable the diagnosis of all rare genetic diseases, Am. J. Hum. Genet. 100 (2017) 695–705.
[3] G. Costain, S. Walker, M. Marano, D. Veenma, M. Snell, M. Curtis, S. Luca, J. Buera, D. Arje, M.S. Reuter, B. Thiruvahindrapuram, B. Trost, W.W.L. Sung, R.K. C. Yuen, D. Chitayat, R. Mendoza-Londono, D.J. Stavropoulos, S.W. Scherer, C.R. Marshall, R.D. Cohn, E. Cohen, J. Orkin, M.S. Meyn, R.Z. Hayeems, Genome sequencing as a diagnostic test in children with unexplained medical complexity, JAMA Netw. Open 3 (2020).
[4] X.F. Li, X.Y. Zhang, S. Zhang, Z.J. Lu, J.Y. Zhang, J.C. Zhou, B.Z. Li, L. Ou, Rare disease awareness and perspectives of physicians in China: a questionnaire-based study, Orphanet J. Rare Dis. 16 (2021).
[5] C.F. Wright, D.R. FitzPatrick, H.V. Firth, Paediatric genomics: diagnosing rare disease in children, Nat. Rev. Genet. 19 (2018) 325.
[6] K. Rahit, M. Tarailo-Graovac, Genetic modifiers and rare mendelian disease, Genes 11 (2020).
[7] C.F. Wright, D.R. FitzPatrick, H.V. Firth, Paediatric genomics: diagnosing rare disease in children (vol 19, pg 325, 2018), Nat. Rev. Genet. 19 (2018), 327-327.
[8] M. Hao, W. Pu, Y. Li, S. Wen, C. Sun, Y. Ma, H. Zheng, X. Chen, J. Tan, G. Zhang, M. Zhang, S. Xu, Y. Wang, H. Li, J. Wang, L. Jin, The HuaBiao project: whole-exome sequencing of 5000 Han Chinese individuals, J Genet Genomics 48 (2021) 1032–1035.
[9] E. Tambuyzer, B. Vandendriessche, C.P. Austin, P.J. Brooks, K. Larsson, K.I. Miller Needleman, J. Valentine, K. Davies, S.C. Groft, R. Preti, T.I. Oprea, M. Prunotto, Therapies for rare diseases: therapeutic modalities, progress and challenges ahead, Nat. Rev. Drug Discov. 19 (2020) 93–111.
[10] T. Suzuki, C. Huang, H. Fujihira, The cytoplasmic peptide:N-glycanase (NGLY1) - structure, expression and cellular functions, Gene 577 (2016) 1–7.
[11] A.C. Need, V. Shashi, Y. Hitomi, K. Schoch, K.V. Shianna, M.T. McDonald, M.H. Meisler, D.B. Goldstein, Clinical application of exome sequencing in undiagnosed genetic conditions, J. Med. Genet. 49 (2012) 353–361.
[12] 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, D.B. Goldstein, Mutations in NGLY1 cause an inherited disorder of the endoplasmic reticulum-associated degradation pathway, Genet. Med. 16 (2014) 751–758.
[13] 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, J.J.M. Jans, Aspartylglycosamine is a biomarker for NGLY1-CDDG, a congenital disorder of deglycosylation, Mol. Genet. Metabol. 127 (2019) 368–372.
[14] K. Abuduxikuer, L. Zou, L. Wang, L. Chen, 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.
[15] H. Hirayama, Y. Tachida, J. Seino, T. Suzuki, A method for assaying peptide: N-glycanase/N-glycanase 1 activities in crude extracts using an N-glycosylated cyclopeptide, Glycobiology 32 (2022) 110–122.
[16] P. Lipinski, A. Bogdanska, A. Rozdzynska-Swiatkowska, A. Wierzbicka-Rucinska, A. Tylki-Szymanska, NGLY1 deficiency: novel patient, review of the literature and diagnostic algorithm, JIMD Rep 51 (2020) 82–88.
[17] I.M. Rios-Flores, M.A. Bonal-Perez, A. Castellanos-Gonzalez, E. Velez-Gomez, A.M. Bertoli-Avella, L. Bobadilla-Morales, C. Pena-Padilla, V. Appendini-Andrade, A. Corona-Rivera, I. Romero-Valenzuela, J.R. Corona-Rivera, Acute liver failure in a male patient with NGLY1-congenital disorder of deglycosylation, Eur. J. Med. Genet. 63 (2020) 103952.
[18] P.L. Hall, C. Lam, J.J. Alexander, G. Asif, G.T. Berry, C. Ferreira, H.H. Freeze, W.A. Gahl, K.K. Nickander, J.D. Sharer, C.M. Watson, L. Wolfe, K.M. Raymond, Urine oligosaccharide screening by MALDI-TOF for the identification of NGLY1 deficiency, Mol. Genet. Metabol. 124 (2018) 82–86.
[19] I. Dabaj, B. Sudrie-Arnaud, F. Lecoquierre, K. Raymond, F. Ducatez, A.M. Guerrot, S. Snanoudj, S. Coutant, P. Saugier-Veber, S. Marret, G. Nicolas, A. Tebani, S. Bekri, NGLY1 deficiency: a rare newly described condition with a typical presentation, Life 11 (2021).
[20] S. Katiyar, T. Suzuki, B.J. Balgobin, W.J. Lennarz, Site-directed mutagenesis study of yeast peptide:N-glycanase. Insight into the reaction mechanism of deglycosylation, J. Biol. Chem. 277 (2002) 12953–12959.
[21] M. Mirdita, S. Ovchinnikov, M. Steinegger, ColabFold: making protein folding accessible to all, Nat Methods 19 (2022) 679–682.
[22] 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, D. Hassabis, Highly accurate protein structure prediction with AlphaFold, Nature 596 (2021) 583–589.
[23] X. Robert, P. Gouet, Deciphering key features in protein structures with the new ENDscript server, Nucleic Acids Res. 42 (2014) W320–W324.
[24] G.M. Sastry, M. Adzhigirey, T. Day, R. Annabhimoju, W. Sherman, Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments, J. Comput. Aided Mol. Des. 27 (2013) 221–234.
[25] T. Suzuki, The cytoplasmic peptide:N-glycanase (Ngly1)-basic science encounters a human genetic disorder, J. Biochem. 157 (2015) 23–34.
[26] R.A. Friesner, R.B. Murphy, M.P. Repasky, L.L. Frye, J.R. Greenwood, T.A. Halgren, P.C. Sanschagrin, D.T. Mainz, Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes, J. Med. Chem. 49 (2006) 6177–6196.
[27] G. Sun, X. Yu, C. Bao, L. Wang, M. Li, J. Gan, D. Qu, J. Ma, L. Chen, Identification and characterization of a novel prokaryotic peptide: N-glycosidase from Elizabethkingia meningoseptica, J. Biol. Chem. 290 (2015) 7452–7462.
[28] W. Zhang, H. Wang, H. Tang, P. Yang, Endoglycosidase-mediated incorporation of 18O into glycans for relative glycan quantitation, Anal. Chem. 83 (2011) 4975–4981.
[29] X. Cao, Q.H. Shang, X.L. Chi, W. Zhang, H.M. Xiao, M.M. Sun, G. Chen, Y. An, C.L. Lv, L. Wang, Y.M. Nan, C.Y. Chen, Z.N. Tan, X.E. Liu, H. Zhuang, Serum N-glycan markers for diagnosing liver fibrosis induced by hepatitis B virus, World J. Gastroenterol. 26 (2020) 1067–1079.
[30] M. Cong, X. Ou, J. Huang, J. Long, T. Li, X. Liu, Y. Wang, X. Wu, J. Zhou, Y. Sun, Q. Shang, G. Chen, H. Ma, W. Xie, H. Piao, Y. Yang, Z. Gao, X. Xu, Z. Tan, C. Chen, N. Zeng, S. Wu, Y. Kong, T. Liu, P. Wang, H. You, J. Jia, H. Zhuang, A predictive model using N-glycan biosignatures for clinical diagnosis of early hepatocellular carcinoma related to hepatitis B virus, OMICS 24 (2020) 415–423.
[31] X.E. Liu, L. Desmyter, C.F. Gao, W. Laroy, S. Dewaele, V. Vanhooren, L. Wang, H. Zhuang, N. Callewaert, C. Libert, R. Contreras, C. Chen, N-glycomic changes in hepatocellular carcinoma patients with liver cirrhosis induced by hepatitis B virus, Hepatology 46 (2007) 1426–1435.
[32] A. Pandey, J.M. Adams, S.Y. Han, H. Jafar-Nejad, NGLY1 deficiency, a congenital disorder of deglycosylation: from disease gene function to pathophysiology, Cells (2022) 11.
[33] X. Miao, J. Wu, H. Chen, G. Lu, Comprehensive analysis of the structure and function of peptide:N-glycanase 1 and relationship with congenital disorder of deglycosylation, Nutrients 14 (2022).
[34] H.R. Lu, K.J. Xiao, Z.X. Tian, Benchmark of site- and structure-specific quantitative tissue N-glycoproteomics for discovery of potential N-glycoprotein markers: a case study of pancreatic cancer, Glycoconj. J. 38 (2021) 213–231.
[35] C. Wang, M. Yang, X. Gao, C. Li, Z. Zou, J. Han, L. Huang, Z. Wang, The ammonia-catalyzed release of glycoprotein N-glycans, Glycoconj. J. 35 (2018) 411–420.
[36] A. Guttman, T. Pritchett, Capillary gel electrophoresis separation of high-mannose type oligosaccharides derivatized by 1-aminopyrene-3,6,8-trisulfonic acid, Electrophoresis 16 (1995) 1906–1911.
[37] M. Hao, W. Pu, Y. Li, S. Wen, C. Sun, Y. Ma, H. Zheng, X. Chen, J. Tan, G. Zhang, M. Zhang, S. Xu, Y. Wang, H. Li, J. Wang, L. Jin, The HuaBiao project: whole-exome sequencing of 5,000 han Chinese individuals, J Genet Genomics 48 (2021) 1032–1035.
[38] N. Pagano, V. Longobardi, C. De Canditiis, C. Zuchegna, A. Romano, K. Michal Andrzej, M.E. Pero, B. Gasparrini, Effect of caspase inhibitor Z-VAD-FMK on bovine sperm cryotolerance, Reprod. Domest. Anim. 55 (2020) 530–536.
[39] A.S. Spirin, V.I. Baranov, L.A. Ryabova, S.Y. Ovodov, Y.B. Alakhov, A continuous cell-free translation system capable of producing polypeptides in high yield, Science 242 (1988) 1162–1164.
[40] A. Zemella, L. Thoring, C. Hoffmeister, S. Kubick, Cell-free protein synthesis: pros and cons of prokaryotic and eukaryotic systems, Chembiochem 16 (2015) 2420–2431.
[41] T. Sawasaki, T. Ogasawara, R. Morishita, Y. Endo, A cell-free protein synthesis system for high-throughput proteomics, Proc. Natl. Acad. Sci. U.S.A. 99 (2002) 14652–14657.
[42] J.A. Rothblatt, D.I. Meyer, Secretion in yeast: reconstitution of the translocation and glycosylation of alpha-factor and invertase in a homologous cell-free system, Cell 44 (1986) 619–628.
[43] K. Noda, Y. Tsuda, F. Kozawa, M. Igarashi, K. Shimizu, J. Arikawa, K. Yoshimatsu, The polarity of an amino acid at position 1891 of severe fever with thrombocytopenia syndrome virus L protein is critical for the polymerase activity, Viruses 13 (2020).
[44] E.V. Pappa, A.A. Zompra, Z. Diamantopoulou, Z. Spyranti, G. Pairas, F.N. Lamari, P. Katsoris, G.A. Spyroulias, P. Cordopatis, Structure-activity studies of lGnRH-III through rational amino acid substitution and NMR conformational studies, Biopolymers 98 (2012) 525–534.
[45] C. Ehre, Z.L. Rushton, B. Wang, L.N. Hothem, C.B. Morrison, N.C. Fontana, M.R. Markovetz, M.F. Delion, T. Kato, D. Villalon, W.R. Thelin, C.R. Esther Jr., D. B. Hill, B.R. Grubb, A. Livraghi-Butrico, S.H. Donaldson, R.C. Boucher, An improved inhaled mucolytic to treat airway muco-obstructive diseases, Am. J. Respir. Crit. Care Med. 199 (2019) 171–180.
[46] A. Minarini, S. Ferrari, M. Galletti, N. Giambalvo, D. Perrone, G. Rioli, G.M. Galeazzi, N-acetylcysteine in the treatment of psychiatric disorders: current status and future prospects, Expet Opin. Drug Metabol. Toxicol. 13 (2017) 279–292.
[47] K. Li, J. Zhang, J. Cao, X. Li, H. Tian, 1,4-Dithiothreitol treatment ameliorates hematopoietic and intestinal injury in irradiated mice: potential application of a treatment for acute radiation syndrome, Int. Immunopharm. 76 (2019) 105913.
[48] H. Fujihira, Y. Masahara-Negishi, M. Tamura, C.C. Huang, Y. Harada, S. Wakana, D. Takakura, N. Kawasaki, N. Taniguchi, G. Kondoh, T. Yamashita, Y. Funakoshi, 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).
[49] A. Galeone, S.Y. Han, C. Huang, A. Hosomi, T. Suzuki, H. Jafar-Nejad, Tissue-specific regulation of BMP signaling by Drosophila N-glycanase 1, Elife 6 (2017).
[50] N. Habibi-Babadi, A. Su, C.E. de Carvalho, A. Colavita, The N-glycanase png-1 acts to limit axon branching during organ formation in Caenorhabditis elegans, J. Neurosci. 30 (2010) 1766–1776.
[51] M. Might, M. Wilsey, The shifting model in clinical diagnostics: how next-generation sequencing and families are altering the way rare diseases are discovered, studied, and treated, Genet. Med. 16 (2014) 736–737.
[52] C. Wang, S. Qiang, W. Jin, X. Song, Y. Zhang, L. Huang, Z. Wang, Reductive chemical release of N-glycans as 1-amino-alditols and subsequent 9-fluorenyl-methyloxycarbonyl labeling for MS and LC/MS analysis, J. Proteonomics 187 (2018) 47–58.
[53] M. Aoyama, N. Hashii, W. Tsukimura, K. Osumi, A. Harazono, M. Tada, M. Kiyoshi, A. Matsuda, A. Ishii-Watabe, Effects of terminal galactose residues in mannose alpha1-6 arm of Fc-glycan on the effector functions of therapeutic monoclonal antibodies, mAbs 11 (2019) 826–836.
[54] A.O. Caglayan, S. Comu, J.F. Baranoski, Y. Parman, H. Kaymakcalan, G.T. Akgumus, C. Caglar, D. Dolen, E.Z. Erson-Omay, A.S. Harmanci, K. Mishra-Gorur, H. H. Freeze, K. Yasuno, K. Bilguvar, M. Gunel, NGLY1 mutation causes neuromotor impairment, intellectual disability, and neuropathy, Eur. J. Med. Genet. 58 (2015) 39–43.
[55] J. Heeley, M. Shinawi, Multi-systemic involvement in NGLY1-related disorder caused by two novel mutations, Am. J. Med. Genet. 167A (2015) 816–820.
[56] D.G. Bosch, F.N. Boonstra, N. de Leeuw, R. Pfundt, W.M. Nillesen, J. de Ligt, C. Gilissen, S. Jhangiani, J.R. Lupski, F.P. Cremers, B.B. de Vries, Novel genetic causes for cerebral visual impairment, Eur. J. Hum. Genet. 24 (2016) 660–665.
[57] 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, L. Wolfe, Prospective phenotyping of NGLY1-CDDG, the first congenital disorder of deglycosylation, Genet. Med. 19 (2017) 160–168.
[58] C.A. Chang, X.C. Wei, S.R. Martin, D.S. Sinasac, W. Al-Hertani, Transiently elevated plasma methionine, S-adenosylmethionine and S-adenosylhomocysteine: unreported laboratory findings in a patient with NGLY1 deficiency, a congenital disorder of deglycosylation, JIMD Rep 49 (2019) 21–29.
[59] B.J. van Keulen, J. Rotteveel, M.J.J. Finken, Unexplained death in patients with NGLY1 mutations may be explained by adrenal insufficiency, Phys. Rep. 7 (2019) e13979.
[60] P. Lipinski, J. Cielecka-Kuszyk, P. Socha, A. Tylki-Szymanska, Liver involvement in NGLY1 congenital disorder of deglycosylation, Pol. J. Pathol. 71 (2020) 66–68.

[61] 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, R.J. Rodenburg, Variants in NGLY1 lead to intellectual disability, myoclonus epilepsy, sensorimotor axonal polyneuropathy and mitochondrial dysfunction, Clin. Genet. 97 (2020) 556–566.
[62] P. Lipari Pinto, C. Machado, P. Janeiro, J. Dupont, S. Quintas, A.B. Sousa, A. Gaspar, NGLY1 deficiency-A rare congenital disorder of deglycosylation, JIMD Rep 53 (2020) 2–9.
[63] H. Ge, Q. Wu, H. Lu, Y. Huang, T. Zhou, D. Tan, ZhongqinJin, Two novel compound heterozygous mutations in NGLY1as a cause of congenital disorder of deglycosylation: a case presentation, BMC Med. Genet. 21 (2020) 135.
[64] A. Kariminejad, M. Shakiba, M. Shams, P. Namiranian, M. Eghbali, S. Talebi, M. Makvand, J. Jaeken, H. Najmabadi, R.C. Hennekam, NGLY1 deficiency: novel variants and literature review, Eur. J. Med. Genet. 64 (2021) 104146.
[65] W. Wang, L. Sun, T. Li, Y. Ma, J. Li, Y. Liu, M. Li, L. Wang, C. Li, Y. Xie, Y. Wen, M. Liang, L. Chen, S. Tong, A human monoclonal antibody against small envelope protein of hepatitis B virus with potent neutralization effect, mAbs 8 (2016) 468–477.