

RESEARCH

Open Access



# The introduction of an N-glycosylation site into prochymosin greatly enhances its production and secretion by *Pichia pastoris*

Nan Wang<sup>†</sup>, Caifeng Yang<sup>†</sup>, Huakang Peng, Wenfang Guo, Mengqi Wang, Gangqiang Li and Dehu Liu<sup>\*</sup>

## Abstract

**Background:** N-glycosylation is one of the most important post-translational modifications. Many studies have shown that N-glycosylation has a significant effect on the secretion level of heterologous glycoproteins in yeast cells. However, there have been few studies reporting a clear and unified explanation for the intracellular mechanism that N-glycosylation affect the secretion of heterologous glycoproteins so far. *Pichia pastoris* is an important microbial cell factory producing heterologous protein. It is of great significance to study the effect of N-glycosylation on the secretion level of heterologous protein. Camel chymosin is a glycoprotein with higher application potential in cheese manufacturing industry. We have expressed camel prochymosin in *P. pastoris* GS115, but the lower secretion level limits its industrial application. This study attempts to increase the secretion level of prochymosin through N-glycosylation, and explore the molecular mechanism of N-glycosylation affecting secretion.

**Results:** Adding an N-glycosylation site at the 34th amino acid of the propeptide of prochymosin significantly increased its secretion in *P. pastoris*. N-glycosylation improved the thermostability of prochymosin without affecting the enzymatic activity. Immunoprecipitation coupled to mass spectrometry (IP-MS) analysis showed that compared with the wild prochymosin (chy), the number of proteins interacting with N-glycosylated mutant (chy34) decreased, and all differential interacting proteins (DIPs) were down-regulated in chy34-GS115 cell. The DIPs in endoplasmic reticulum were mainly concentrated in the misfolded protein pathway. Among the five DIPs in this pathway, overexpression of BiP significantly increased the secretion of chy. The knockout of the possible misfolded protein recognition elements, UDP-glycose:glycoprotein glucosyltransferase 1 and 2 (UGGT1/2) had no effect on the growth of yeast cells and the secretion of prochymosin.

**Conclusions:** In conclusion, N-glycosylation increased the secretion of prochymosin in *P. pastoris* through the adjustment of intracellular interacted proteins. The results of our study may help to elucidate the molecular mechanism of N-glycosylation affecting secretion and provide a new research method to improve the secretion of heterologous glycoprotein in *P. pastoris*.

**Keywords:** N-glycosylation, Prochymosin, Secretion, *Pichia pastoris*, Differential interacting proteins

## Background

In eukaryotic cells, N-glycosylation is one of the most important post-translational modifications and plays an important role in biological processes [1]. N-glycosylation affects biological activity, thermostability, protease tolerance, signal transduction, localization, and the secretion of glycoproteins [2–9]. N-glycosylation of eukaryotic

<sup>†</sup>Nan Wang and Caifeng Yang contributed equally to this work

\*Correspondence: liudehu2006@126.com

Biotechnology Research Institute, Chinese Academy of Agricultural Sciences, Beijing 100081, China



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

cells is a conservative process. In the endoplasmic reticulum (ER), oligosaccharide transferase (OST), an ER membrane-localized complex, transfers an oligosaccharide (Glc3Man9GlcNAc2) to the Asn residue within the Asn-Xxx-Ser/Thr (where Xxx is any amino acid except Pro) recognition sequence in a nascent protein. Then, the outer three glucose molecules of the oligosaccharide are successively cleaved by  $\alpha$ -glucosidase I and II. If the protein folding is completed quickly, it will be transported to the Golgi apparatus via vesicular trafficking for further glycosylation [10, 11], and finally, the glycoproteins are localized to a certain organelle or secreted to the outside of the cell. In contrast, if the glycoproteins does not fold properly, many studies inferred that the UDP-glucose: glycoprotein glucosyltransferase (UGGTs, including UGGT1 and UGGT2) will recognize it and adds a single glucose to the N-glycan [12, 13]. The GlcMan9GlcNAc2 acts as a signal to recruit the lectins calnexin (CNX) and other molecular chaperones, like Protein disulfide isomerase (PDI), to help the glycoproteins complete the folding, while preventing the non-folded glycoproteins from leaving the ER. For glycoproteins that cannot complete the correct folding over a long period of time,  $\alpha$ -mannosidase will remove mannose and the remaining N-glycan as a signal to enter the degradation pathway. This process is call ER-associated degradation (ERAD) [14]. Meanwhile, the BiP, a molecular chaperone in the ER lumen, has been shown to interacts with unfolded proteins and participants in the ERAD [15, 16]. In the glycoprotein processing, it is worth noting that the function of UGGTs in the important microbial cell factory *Pichia pastoris* (*P. pastoris*) has not been confirmed.

Many studies have shown that N-glycosylation has a significant effect on the secretion level of heterologous glycoproteins in yeast cells [6, 7]. The position of glycosylation in glycoprotein plays a key role in this process [17]. Most N-glycosylation can improve secretion of glycoproteins [18–20], but there are also cases where the secretion is reduced [21] or not changed [22, 23]. However, there have been no studies reporting a clear analysis of and unified explanation for the intracellular mechanism of N-glycosylation affecting the secretion of heterologous glycoproteins. *P. pastoris* is an important microbial cell factory producing heterologous protein. It is of great significance to study the effect of N-glycosylation on the secretion level of heterologous glycoprotein.

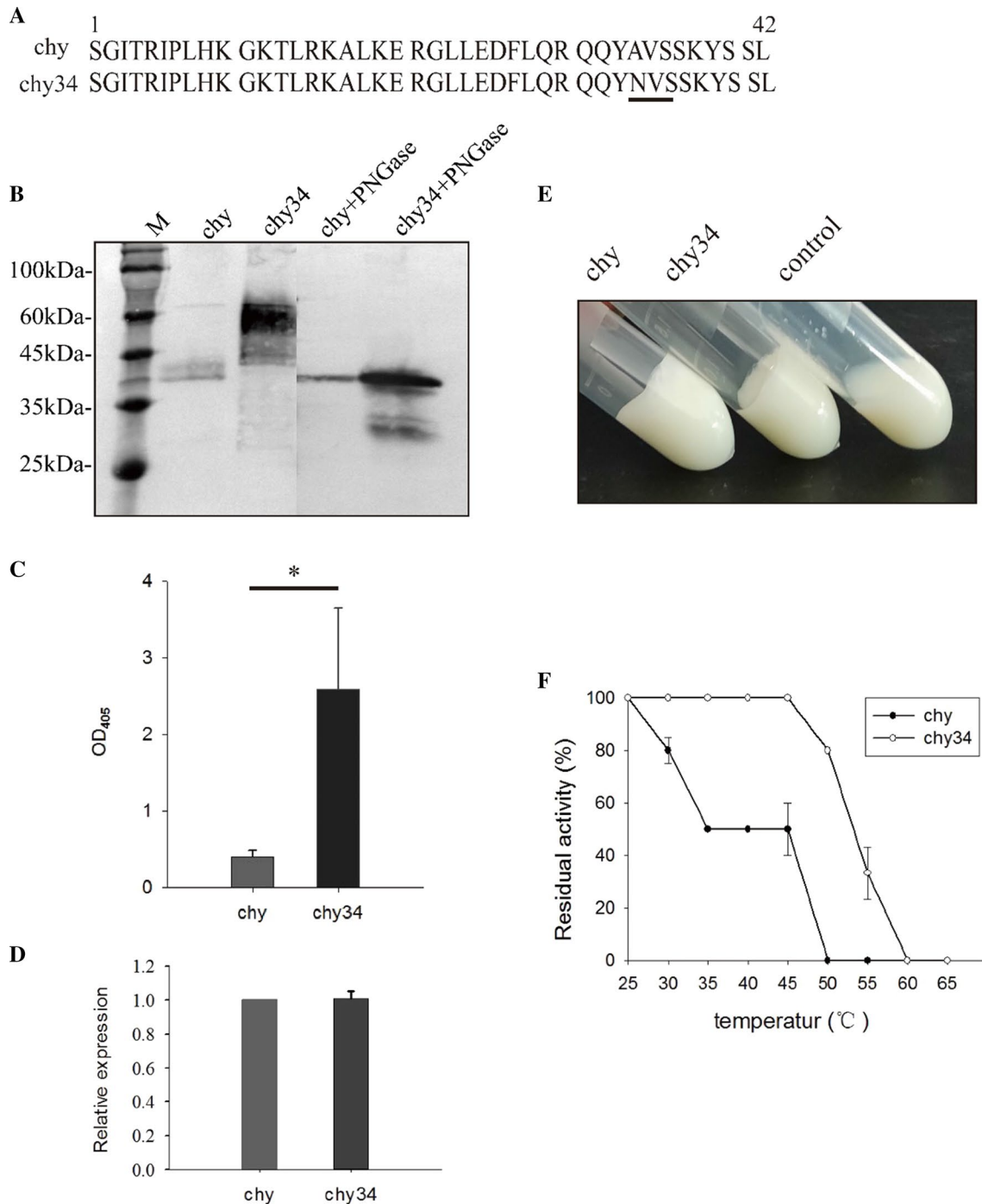
Chymosin is an important enzyme in cheese manufacturing industry. Compared with widely used bovine chymosin, camel chymosin has higher clotting activity, specificity of protease activity, and thermostability [24, 25]. These characteristics make camel chymosin more advantageous in reducing the cost of production, transportation, storage and improving the flavor of cheese.

The inactive camel prochymosin contains 365 amino acid residues. Its N-terminal has a propeptide of 42 residues, which is remove in an acidic environment, thereby leading to activation of the enzyme, this process known as autocatalytic cleavage. In our previous study, recombinant *P. pastoris* GS115 host expressing camel prochymosin was constructed [26], and the secretion level was approximately 37 mg/L under shaking flask conditions. This secreted level is too low to meet the needs of industrial application. The purpose of this study is to increase the secretion of prochymosin in *P. pastoris* as well as explore the intracellular molecular mechanism of N-glycosylation affecting secretion.

## Results

### N-glycosylation improved thermostability of prochymosin and increased secretion in *P. pastoris*

In order to increase the secretion of prochymosin in *P. pastoris* without affecting the clotting activity, an N-glycosylated conserved sequence at amino acid position 34–36 of the propeptide (Fig. 1A) was constructed by site-directed mutation (A34N). The mutant prochymosin (chy34) could be secrete in *P. pastoris*, and its molecular weight becomes larger (from ~40 kDa to ~60 kDa). After treatment with PNGase F, the molecular weight of chy34 decreased to the same as that of wild prochymosin (chy), indicating that the glycosylation site was N-glycosylated (Fig. 1B). The secretion level of chy34 was significantly higher than that of chy (Fig. 1C). There was no significant difference in mRNA level (Fig. 1D), indicating that the difference in protein secretion was not caused by copy number. The enzymatic activity analysis showed that the N-glycosylation mutations in the propeptide region did not affect its autocatalytic cleavage activity (Fig. 1E). Chy and chy34 were incubated at different temperatures. After 8 h of incubation at 45 °C, the residual activity of chy decreased to about 50%, while chy34 remained fully active (Fig. 1F), indicating that N-glycosylation improved the thermostability of the prochymosin. This improvement of thermostability was further verified on the secondary and tertiary structure of protein by circular dichroism spectroscopy (CD) and fluorescence spectroscopy, respectively. The CD spectra (Additional file 1: Fig. S5A) revealed that there was no significant difference in the proportion of secondary structure between the chy and chy34 proteins. Fluorescence emission spectra (Additional file 1: Fig. S5B) showed that the maximum emission of chy34 had a slight blue shift, and the fluorescence density decreased, it indicated that the thermostability of chy34 was improved, which may be attributed to the effect of N-glycosylation on the tertiary structure of protein.



**Fig. 1** The effect of N-glycosylation on molecular weight, autocatalytic cleavage activity, thermal stability of prochymosin and the secretion in *P. pastoris* GS115. **A** Mutant site of N-glycosylation in the propeptide of prochymosin, conserved N-glycosylated sequences were indicated by underlines; **B** molecular weight and de-glycosylation analysis of chy and chy34 by Western-blot; **C** Determination of the secretion of chy and chy34 by ELISA; **D** The mRNA levels of chy and chy34 in host cells determined by qRT-PCR; **E** Analysis of autocatalytic cleavage of the propeptide by confirming the milk-clotting acidity, The control was the supernatant of GS115 culture carried the pPIC9K vector; **F** Thermostability of chy and chy34 when stored at different temperatures for 8 h prior to enzymatic activity assay. Error bars indicate the standard deviation of three independent experiments, student's t-test was used to compare difference between two groups, \*P < 0.05

### Identification of differential proteins interacting with chy and chy34

In order to explore the molecular mechanism of N-glycosylation affecting the secretion of prochymosin, the differential interacting proteins (DIPs) with chy and chy34 in two host cells were determined by immunoprecipitation coupled to mass spectrometry (IP-MS) analysis. First, the agarose beads combined with anti-prochymosin antibody were used to immunize and precipitate the proteins that directly or indirectly interacting with chy and chy34. Then two groups of proteins were analyzed by mass spectrometry. The results of mass spectrometry searched against the UniProt database. The DIPs were identified by comparing all the proteins obtained from the database. Then GO annotation, KEGG enrichment analysis and COG analysis were performed on the DIPs.

After the proteins sample obtained by immunoprecipitation were digested by trypsin, the peptide length is mainly between 8 and 20 (Fig. S1A), which conforms to the regular of trypsin digestion and meets the standard for mass spectrometry. By the detection of mass spectrometry, 147 proteins interacting with chy34 and 339 proteins interacting with chy were identified (Fig. 2A), and the number of spectra, peptides and proteins interacting with chy was higher than that of chy34 (Additional file 1: Fig. S1B). Taking the fold change  $\geq 1.5$  as the screening standard, 124 DIPs were obtained (Additional file 2: Table S1). All DIPs were down-regulated in chy34 (Additional file 1: Fig. S1C), which showed higher secretion in *P. pastoris*. Among the DIPs, 87 interacted with both chy and chy34, and 37 interacted only with chy (Fig. 2B). These results showed that the type and number of proteins interacting with chy were significantly more than those of chy34.

### Annotation and enrichment analyses of DIPs

GO annotation and enrichment analysis was performed on 124 DIPs. The top terms (P Value  $< 0.05$ ) were ranked based on the P value. The top Biological Process (BP) terms were organonitrogen compound metabolic process, protein metabolic process, organic substance biosynthetic process, etc.; the top Cellular Component (CC) terms were cytosolic ribosome, ribosomal subunit, pre-ribosome, etc.; the top Molecular Function (MF) terms were ribonucleotide binding, nucleoside phosphate binding, nucleotide binding etc. (Fig. 2C). The KEGG pathway enrichment analysis (Fig. 2D) showed that the DIPs were mostly involved in Ribosome, Glycolysis/Gluconeogenesis, Autophagy—yeast, Protein processing in endoplasmic reticulum, etc. COG analysis (Fig. 2E) showed that the functions of DIPs were enriched in Translation, ribosomal structure and biogenesis, Energy production and

conversion, Posttranslational modification, protein turnover, chaperones, etc. From the annotation and enrichment analysis results of DIPs, it can be seen that these DIPs are mainly related to protein synthesis, metabolism, energy production/transformation and homeostasis maintenance.

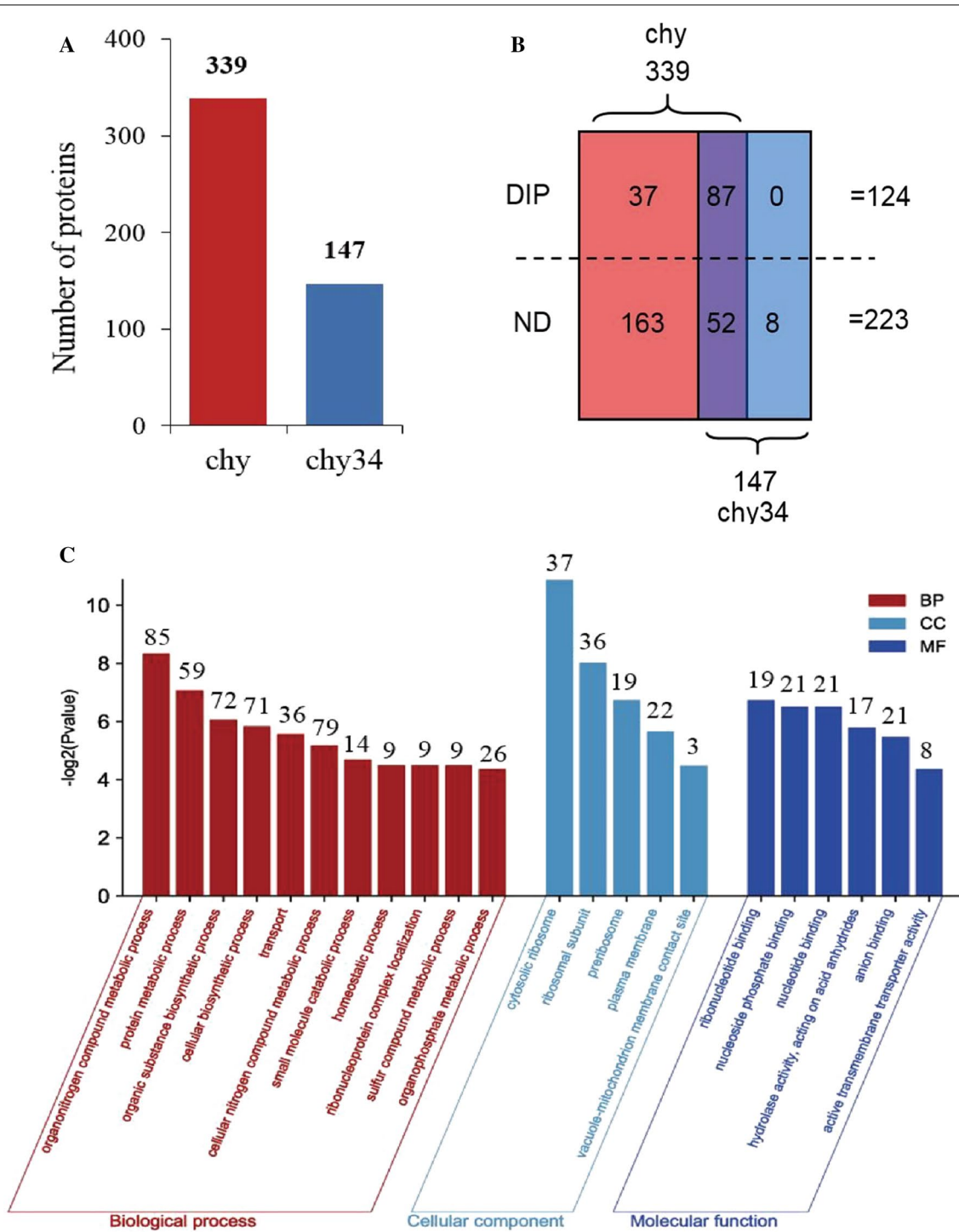
The translation, folding and modification of N-glycosylated proteins begin in the endoplasmic reticulum (ER). In the enriched KEGG pathways, there are six proteins in the ER related “protein processing in endoplasmic reticulum” pathway: PDI (Gene symbol: PAS\_chr1-1\_0160), Hsp90 (Gene symbol: PAS\_chr1-4\_0130), BiP (Gene symbol: PAS\_chr2-1\_0140), Hsp70 (Gene symbol: PAS\_chr3\_0230), Hsp70 (Gene symbol: PAS\_chr3\_0731), Hsp70 (Gene symbol: PAS\_chr4\_0552). These proteins are mainly involved in the metabolic pathway when glycoproteins misfolded (Fig. 3). The red rectangles in Fig. 3 mean up-regulated molecular chaperone or enzyme in chy sample. This indicated that in the ER, more molecular chaperones and enzymes were involved in the folding, secretion or degradation of wild chy. It also implied that after N-glycosylation of prochymosin, the proteins interacting with prochymosin in ER were down-regulated and the glycoprotein was released into the secretory pathway, that is, the added N-glycosylation changed the metabolic pathway of prochymosin in ER.

### BiP overexpression increased the secretion of chy

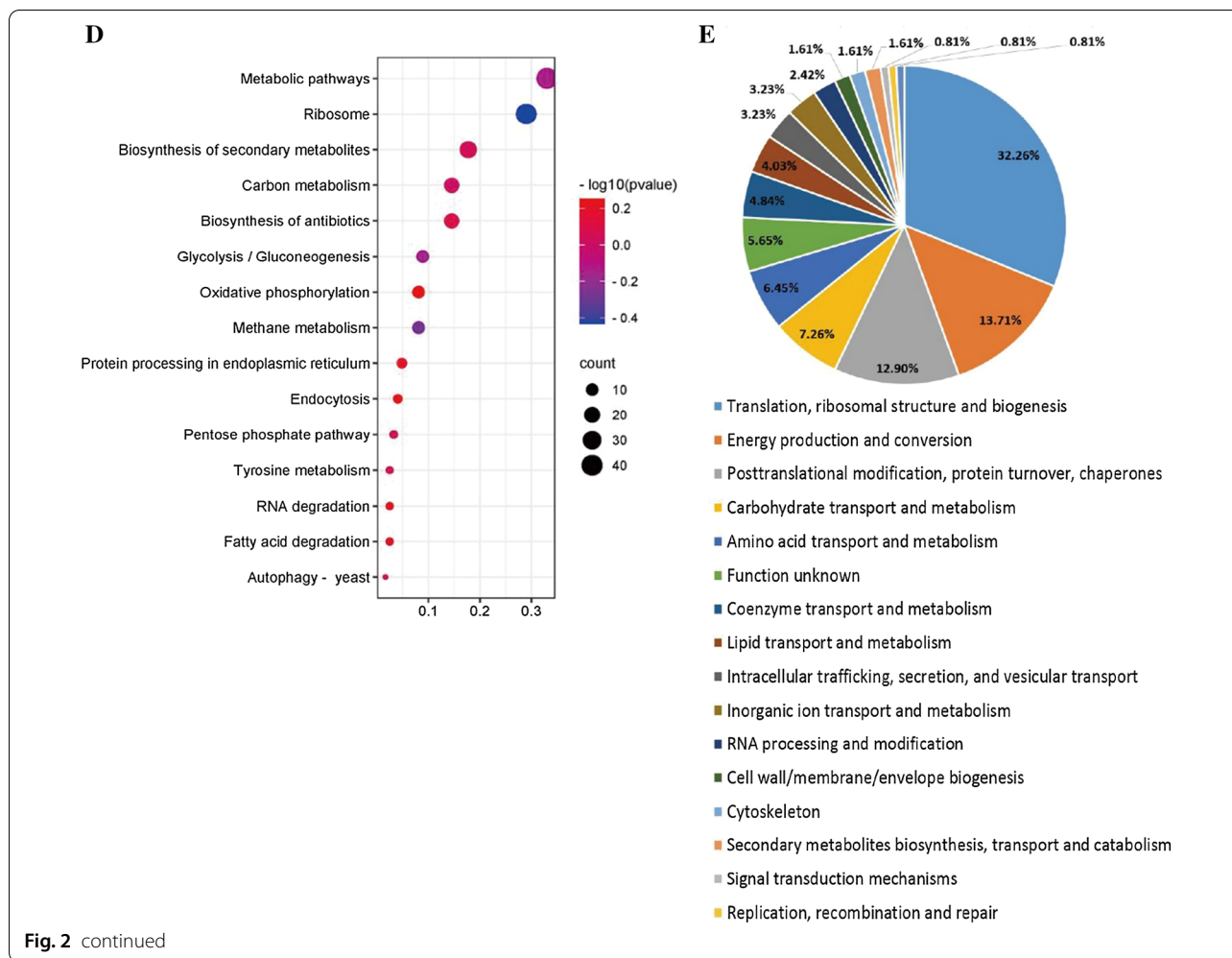
In the ER related “protein processing in ER” pathway (Fig. 3), the proteins interacting with chy are significantly up-regulated, we speculate that these proteins may affect the secretion of chy. Therefore, five up-regulated proteins PDI and BiP in ER, PEP4 (Gene symbol: pas\_chr3\_1087), PRB1 (Gene symbol: pas\_chr1-1\_0226) in autophagy-yeast pathway (ppa04138) and ERAD related protein Hsp70 (Gene symbol: pas\_chr3\_0230) were selected for co-expression to study whether these proteins has an effect on the secretion of chy. The results showed that BiP increased the secretion of chy. Other four proteins had no effect. All six proteins had no significant effect on the secretion of chy34 (Fig. 4A, B).

### The knockout of UGGTs did not affect the secretion of chy and chy34

The above studies show that N-glycosylation changes the interacting proteins of prochymosin in ER. We speculated that there are elements that recognize N-glycosylation modification in ER. Previous studies have shown that, in other species, UGGTs (including two isoforms: UGGT1 and UGGT2) can recognize the folding state of glycoproteins [12, 13]. However, in this study, no UGGTs proteins were detected in the mass spectrometry results of the two samples.



**Fig. 2** Bioinformatics analysis of IP-MS data. **A** Number of differential interacting proteins (DIPs); **B** Distribution of DIPs and no differentially interacting proteins (ND) in chy and chy34 samples; **C** GO function classification of DIPs; **D** KEGG enrichment analysis of DIPs; **E** COG analysis of DIPs. The terms of p Value < 0.05 were displayed



This suggests that the UGGTs may not be the recognition element of prochymosin protein folding, or there may be other intermediate signal molecules. In order to study the role of UGGTs in prochymosin secretion pathway, we knocked out UGGT1 and UGGT2 in host cells (Additional file 1: Fig. S3). The knockout of UGGT1 and UGGT2 did not affect the growth of yeast cells (Additional file 1: Fig. S4). The genes of *chy* and *chy34* were transformed into GS115 ( $\Delta$ UGGTs) strain, and their secretion levels did not change significantly (Fig. 4C).

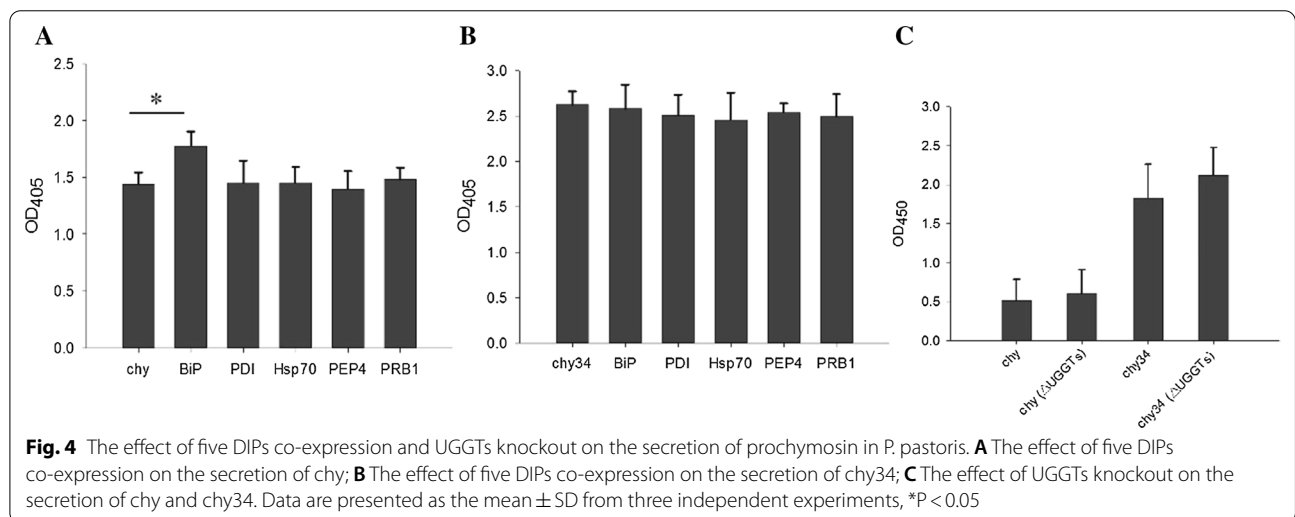
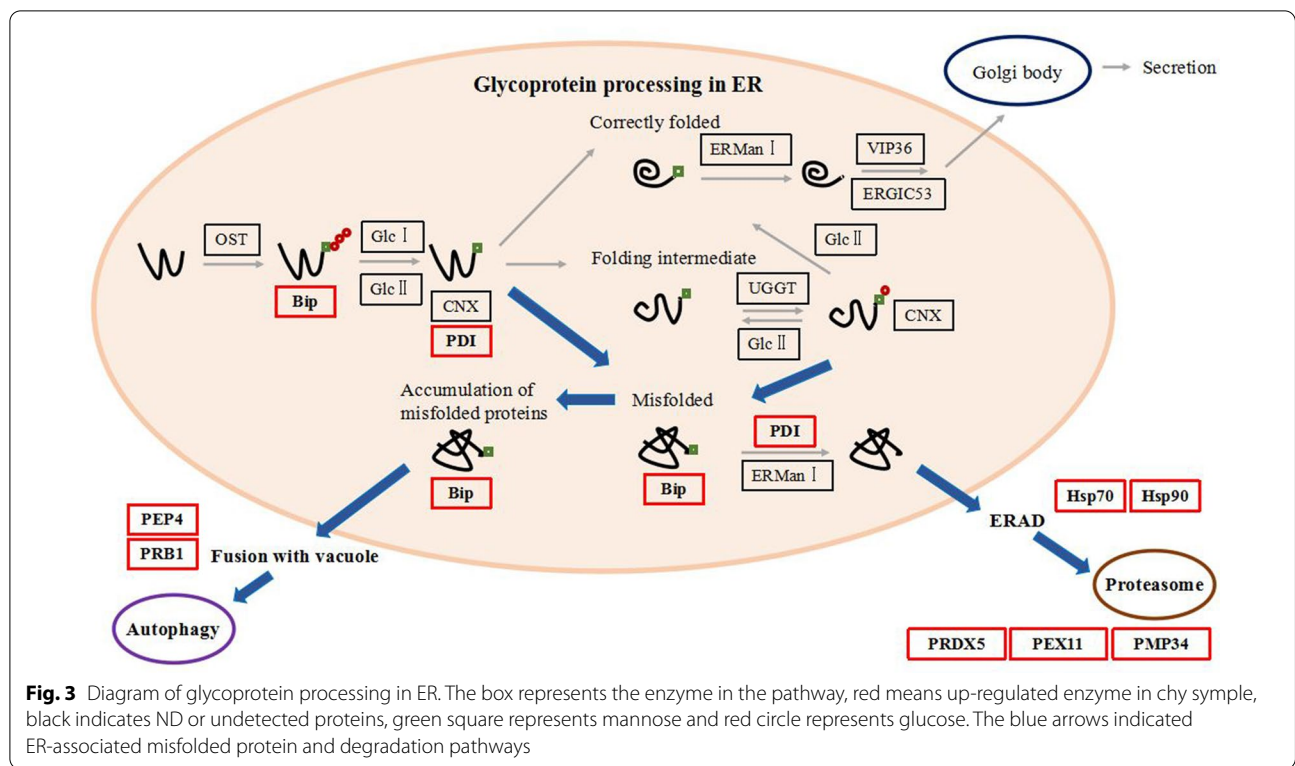
## Discussion

Nowadays, many eukaryotic expression systems have been commercialized [37], which including monoclonal antibodies production in mammalian cells [38], some viral vaccines production in baculovirus expression vector system [39], hepatitis B vaccine production in *Saccharomyces cerevisiae* [40] and insulin glargine production in *P. pastoris* [37], all these expression systems have the ability of N-glycosylation modification. More than half of

all eukaryotic proteins are glycoproteins, and about 90% of them are N-glycosylated [10]. N-glycosylation exerts many biological functions in organisms. Among them, increasing glycoprotein secretion is one of the important functions. However, as the studies on the mechanisms of N-glycosylation affecting glycoproteins secretion, specially the bioinformatics analysis of intracellular relevant proteins were fewer, there are no reports on the application of N-glycosylation increasing protein expression to the industrial reality so far. Present study provided a preliminary research basis for the above question.

## N-glycosylation affects prochymosin protein properties

This study found that N-glycosylation increased the secretion of prochymosin, and the molecular weight and thermostability of the protein also increased significantly. Petrescu et al. reported that, oligosaccharide chains can reduce the conformational freedom of the Asn residue side chain and increase the structural rigidity of nearby amino acids, making them less susceptible to thermal



denaturation [28]. Additionally, the key of the influence of oligosaccharide chain on the stability of protein structure lies in the N-glycosylated position, not the length of the oligosaccharide chain. At present, the main view is the oligosaccharide chains are hydrophilic macromolecules, and their presence makes it more difficult for glycoprotein to form a precipitate in an aqueous solution. Our research is consistent with these views, N-glycosylation

significantly improved the thermostability of prochymosin. However, not all N-glycosylation of proteins can improve secretion, it may be that these N-glycosylation positions have no significant effect on thermostability, or only were weak glycosylated [29] or oligosaccharide chain was located inside the protein molecule. More experimental data are needed to explain these opposite phenomenon. On the other hand, Western-blot results

(Fig. 1B) showed that after N-glycosylation at position 34, the molecular weight increased significantly, indicating that the protein was hyper-glycosylated. The hyper-glycosylation process occurred in the Golgi apparatus [30]. However, it has been found that the Golgi apparatus does not determine the secretion of glycoproteins [31]. Therefore, ER is the key organelle affecting prochymosin secretion.

#### **N-glycosylation changes intracellular proteins interacting with prochymosin**

IP-MS results showed that the proteins interacting with chy were significantly more than that with chy34 (Fig. 2A, B), and all DIPs was up-regulated in chy sample. According to the results of GO, KEGG and COG (Fig. 2C, D, E and Additional file 3: Table S2), DIPs are involved in protein synthesis, metabolism, energy production/transformation and homeostasis maintenance, it indicated that the secretion of chy needs more molecular chaperones, consumes more energy to maintain cell homeostasis. Due to its low thermostability, chy binds to BIP and PDI in ER, preventing it from entering the secretory pathway. We measured the intracellular proteins of the two host cells by Western-blot, and did not detect prochymosin (the results were not shown), and this implying that chy did not stay in ER, but quickly entered the degradation pathway. This phenomenon was also verified by the significant up-regulation of the interacting proteins (Hsp70, PEP4 and PRB1) that related to the degradation pathway. When N-glycosylation sites were added to the prochymosin, the interacting protein was down regulated. In ER, prochymosin was separated from BIP and PDI etc., and then entered the secretory pathway. Among the five DIPs derived from ER, yeast-autophagy and ERAD pathway, BiP can increase the secretion of chy, indicating that BiP plays an important role in the mediation of prochymosin metabolic pathway. BiP had no effect on the secretion of chy34, which also verified that BiP selectively interacted with chy. In ER, BiP can bind to newly synthesized polypeptides [32], prevent their aggregation and precipitation, also prevent misfolded polypeptides from leaving the ER and guide them into the degradation pathway [15, 16]. It is further inferred that chy is not easy to form correct folding in ER, but after added an N-glycosylation, it promoted protein folding and increased secretion.

#### **UGGTs is not a recognition element of the quality control system for prochymosin secretion**

In many species, UGGTs is a folding sensor of glycoprotein quality control system in the ER [33–35]. However, in *P. pastoris*, there has been no any reports about UGGTs function identification. In order to explore the effect of UGGTs on the secretion of prochymosin, we

knocked out UGGT1 and UGGT2 of *P. pastoris* genome. The knockout of UGGTs did not affect the growth of yeast cells, indicating that UGGTs is not necessary for cell growth under normal conditions. And then we determined the effect of UGGTs knockout on the secretion of chy and chy34. Surprisingly, UGGTs had no effect on the secretion of both glycoproteins. It indicates that in *P. pastoris*, N-glycosylation can significantly change the secretion of prochymosin even without UGGTs. So we inferred that UGGTs is not the key recognition element of glycoproteins in *P. pastoris*. In *S. cerevisiae*, there is no UGGTs gene in the genome [36], and there may be other components in its glycoprotein quality control system to regulate this pathway. Although there are UGGTs gene in *P. pastoris* genome, our study did not identify its key function on glycoprotein secretion. In *P. pastoris*, whether BiP is the sensor for glycoprotein quality control needs to be further determined. In addition, it is worth noting that this study is limited to the prochymosin, and whether other glycoproteins meet the findings of this study remains to be verified.

#### **Conclusion**

In conclusion, N-glycosylation significantly increased the secretion of prochymosin in *P. pastoris*, and DIPs before and after adding N-glycosylation site were identified by IP-MS. The up-regulated DIPs interacting with chy mainly involved in the misfolded protein pathway. At the same time, it was also found that BiP can improve the secretion of chy, and UGGTs do not play a key role in the secretion of prochymosin. This study is conducive to promote the industrial application of chymosin and provide a research basis for the interpretation of the molecular mechanism of N-glycosylation affecting glycoprotein secretion.

#### **Materials and methods**

##### **Strains, vectors, and reagents**

The yeast expression vector pPIC9K and the yeast strain *P. pastoris* GS115 were obtained from Life Technologies, Inc. (Grand Island, NY, USA). *E. coli* competent cells were purchased from TransGen Biotech (Beijing, China). The restriction enzyme and Peptide N-Glycosidase F (PNGase F) were purchased from New England Biolabs (Ipswich, MA, USA). Skimmed milk powder was purchased from Sigma-Aldrich (St. Louis, MO, USA). All other chemicals and reagents, unless otherwise stated, were acquired from Sinopharm Chemical Reagent Beijing Co., Ltd. (Beijing, China).

##### **Construction of mutant prochymosin expression vector**

The expression vector pPIC9k-chy with the wild prochymosin genetic sequence [7] was used as a polymerase



chain reaction (PCR) template for site-directed mutagenesis. To construct the prochymosin mutant in which one conserved N-glycosylated site, Asn-Val-Ser, was introduced into the propeptide, the codon for 34th Asn replaced the Ala (Fig. 1A). The primers used in site-directed mutagenesis were listed in Table 1. The PCR products were cloned into the *P. pastoris* vector pPIC9K using a Seamless Assembly Cloning Kit (CloneSmarter, USA) to yield the expression vectors carrying mutant prochymosin genes. Wild (*chy*) and mutant (*chy34*) recombinant plasmids were linearized by *Sal* I and then electrotransformed into GS115 by a Bio-Rad Micropulser Electroporator (Bio-Rad, Hercules, CA, USA). Electrotransformation was performed according to the manufacturer's protocol (Life Technologies, USA).

#### Secretion and identification of prochymosin in *P. pastoris*

The positive clones were inoculated in 10 mL of BMGY (1% yeast extract, 2% peptone, and 4% glycerol) medium containing gradient concentration of G418 (0.1, 0.2, 0.4, and 0.8 mg/mL) and cultured at 28 °C and 200 rpm for 48 h. Fourteen clones that could resist 0.4 mg/mL of G418 for each transformation event were selected and cultured at 28 °C and 200 rpm/min in 10 mL of BMGY medium until the OD<sub>600</sub> reached 2.0. The cells were collected and resuspended in 10 mL of BMMY (1% yeast extract, 2% peptone, 0.5% methanol, and 10 mM potassium phosphate buffer at pH 6.0) medium and then cultured at 28 °C and 200 rpm for 72 h. Moreover, 50 µL of 100% methanol were added every 24 h. The cultures were centrifuged at 10,000 rpm for 5 min, and the supernatants containing secreted proteins were collected. Then, SDS-PAGE or Western-blot were carried out to verify the

proteins. ELISA was used to analyze the relative concentrations of *chy* and *chy34* proteins. The methods of SDS-PAGE, Western-blot and Enzyme linked immunosorbent assay (ELISA) were described in [26]. The primary antibody was rabbit anti-prochymosin polyclonal antibody (prepared by Abclonal Technology, Wuhan, China) and the second antibody was goat anti-rabbit IgG conjugated to Alkaline Phosphatase (ab97072, Abcam, Cambridge, MA, USA). During the ELISA reaction, Alkaline Phosphatase reacted with the substrate BCIP/NBT, and the resulting compound has a maximum absorbance at 405 nm. The concentration difference between *chy* and *chy34* samples was directly reflected by the OD<sub>405</sub>. The pre-test results showed that there was no significant difference in the ELISA and Western-blot of 14 samples (the results were not shown). Three samples were randomly selected from 14 samples came from each transformation event to calculated the standard deviation.

#### Real-time quantitative PCR

Two milliliters of BMGY culture (OD<sub>600</sub> was about 2.0) were used to collect cells by centrifugation at 10,000 rpm for 5 min. The total RNA extraction was performed according to the manufacturer's guidelines of the RNA Purification Kit (Yuanpinghao Bio, Beijing, China). The cDNA synthesis was performed using the cDNA Synthesis SuperMix Kit (TransGen Biotech, Beijing, China). The quantitative real-time PCR (qRT-PCR) was performed using a Green qRT-PCR SuperMix Kit (TransGen Biotech, Beijing, China) on a QuantStudio 3 instrument (Thermo Fisher Scientific, USA). The actin gene of *P. pastoris* GS115 was the reference gene. The qRT-PCR primer pair were listed in Table 1. The relative expression

**Table 1** Primers for the site mutation, qPCR and verified of UGGTs gene knockout

Primer name <sup>a</sup>	Sequence (5' to 3')
<i>chy</i> -F	gggtatctctcgagaaaagaccatcaccatcaccactccgggtaccagaatcc
<i>chy34</i> -F	cttgcaagacaacaatacaacgtctcctcaagactcctcttggtgaaggtcgc
<i>chy34</i> -R	gcgaccttaccgaaggaggagtacttgaggagacgttgattgttctttgcaag
<i>chy</i> -R	gtctaaggcgaattaattcggccgcttagatggccttgccaaccgac
qPCR-F	ttgagatccatgccaaccgt
qPCR-R	acgtcaccgaagatccaca
act-F	agtggtcccatcggtcgtag
act-R	gggtgtggtccagatcttt
UGGT1-F	tggcgggattcgtcaagat
UGGT1-R	aatgtgagcagcaaacgcc
UGGT2-F	gccaaactggtcaaacccc
UGGT2-R	tttgactgcgcaactacagc

<sup>a</sup> The primers *chy*-F, *chy34*-F, *chy34*-R, *chy*-R were used to amplified the gene of *chy34*. The primer pair qPCR-F and qPCR-R were used to qRT-PCR of prochymosin transcription in both host cells *chy*-GS115 and *chy34*-GS115, the act-F and act-R as the primer pair of internal reference gene actin. The primer pair UGGT1-F/UGGT1-R and UGGT1-F/UGGT1-R were used to verify the knockout of UGGT1 and UGGT2, respectively

at the mRNA level were calculated by the relative  $2^{-\Delta\Delta CT}$  method [41]. The  $\Delta CT$  value of chy was taken as the control and calculated the mRNA relative expression of chy34.

#### **Enzymatic activity and thermostability assay of prochymosin proteins**

The full-length prochymosin has no enzymatic activity. Only after the autocatalytic cleavage of the propeptide was completed, the mature chymosin can exert its clotting activity. Acidic conditions were required for the autocatalytic cleavage. In order to determine whether the N-glycosylated mutation of the propeptide affected its autocatalytic cleavage, the enzymatic activity was determined after acid-neutralization of the supernatant of the recombinant protein cultures. The enzymatic activity was analyzed as described previously with minor modifications [26]. Substrate was prepared by mixing skim milk powder (26% w/v) in 10 mmol/L of potassium phosphate buffer (pH 5.5). The chy and chy34 expression strains were induced with 0.5% methanol for 72 h, and the supernatants of the cultures were collected by centrifuged at 10,000 rpm for 5 min. Autocatalytic cleavage was performed at pH 2.0 by the addition of 1 M HCl and incubated at 25 °C for 2 h. Then the activated enzyme solution was adjusted the pH to 5.5 with 1 M NaOH. The enzyme reactions were carried out in 2 ml centrifuge tubes with 10  $\mu$ L of enzyme solution and 490  $\mu$ L of substrate at 37 °C. The milk clots were visualized by tilting the tubes.

The thermostability of prochymosin were assessed as described previously with minor modifications [26]. The enzyme solution was first incubated for 8 h at a range of temperatures (from 25 °C to 65 °C), and then the enzymatic activity was measured by the method mentioned above. An average value of three relative activities was taken for each sample.

#### **Determination of the structural characterization of prochymosin protein molecular**

Far-ultraviolet (UV) circular dichroism (CD) spectroscopy was used to analyze the secondary structure fractions of chy and chy34. The supernatant of the culture was filtered by sterile filter unit with 0.45  $\mu$ m PVDF membrane (Merck Millipore). Protein was purified by size exclusion chromatography with sephadex G75 column. SDS-PAGE and Western-blot determined the target proteins. The purified protein was dialyzed in 0.2 mg/mL in 20 mM sodium phosphate buffer (pH 7.4) and concentrated by ultrafiltration membrane (Amicon Ultra-15 Centrifugal Filter Unit, Merck-Millipore). Far-UV CD spectra were taken on a Chirascan Plus (Applied Photophysics) with the cell of 0.5 mm

pathlength. Measurement was performed at 25 °C in the range from 190 to 260 nm. The sodium phosphate buffer was used as blank control. An average of three times scans was taken for each sample. Pro-Data Viewer and CDNN software was used to analyze the CD spectra result and calculate secondary structure fractions.

Hitachi F-7000 fluorescence spectrophotometer (Hitachi) was used to analyze the change of tertiary structure of proteins. Hydrophobic tryptophan was the intrinsic fluorescent group of protein. The tryptophan much more tends to be buried in the inside hydrophobic area of protein molecule, the more the maximum emission peak of the protein tends to blue shift and the lower the fluorescence density, which indicates that the protein molecule has more thermostability [23]. The purified protein sample of 0.1 mg/mL was taken for the fluorescence spectra and the data were collected by excitation at 280 nm and had an emission range of 280–400 nm. Slit widths were set at 5 nm with a scan speed of 12,000 nm/minute.

#### **Immunoprecipitation and label-free quantification mass spectrometry analysis**

After 24 h of methanol induction, the cells were collected by centrifuge at 10,000 rpm for 5 min. The cell lysate and immunoprecipitation were performed according to the manufacturer's protocol of Immunoprecipitation Kit (Sangon Biotech, Shanghai, China). The antibody used for immunoprecipitation was rabbit anti-prochymosin polyclonal antibody (prepared by Abclonal Technology, Wuhan, China). After immunoprecipitation, the proteins were digested by trypsin. Samples from three independent biological repetitions were mixed and used for mass spectrometry analysis. Equal amounts of digested chy and chy34 sample were separated by a liquid chromatography UltiMate 3000 RSLCnano (Thermo Scientific) with capture column (3  $\mu$ m, 120 Å, 100  $\mu$ m  $\times$  20 mm) and analytical column (2  $\mu$ m, 120 Å, 75  $\mu$ m  $\times$  150 mm). The flow rate of the liquid phase is set to 300 nL/min using a gradient buffer formed by Solvent A (3% DMSO, 0.1% formic acid, 97% H<sub>2</sub>O) and Solvent B (3% DMSO, 0.1% formic acid, 97% ACN). Separated peptide fragments were identified using a Q Exactive Plus spectrometer (Thermo Scientific). Each scan cycle consists of one MS full scan (R = 70 K, AGC = 3e6, max IT = 20 ms, scan range = 350–1800 1 m/z) and followed by 15 MS / MS scans (R = 17.5 K, AGC = 2e5, max IT = 100 ms). The MS/MS spectra was searched against the UniProt database (<https://www.uniprot.org>) by use of ProteinPilot (V4.5) with Unused  $\geq$  1.3 for protein identification.

## Bioinformatics

The difference interacted proteins (DIPs) were screened based on the number of spectra of each protein. Calculate the ratio and the average value of the protein spectra (MeanSP) in different samples,  $x = \log_2$  ratio,  $y = \log_2$  MeanSP, The DIPs screening boundary line  $y = C / (x - \times 0)$ . Proteins distributed outside the boundary line were labeled as DIPs. DIPs with spectra difference more than 1.5-fold were labeled “±” (the boundary lines were  $y_1$  and  $y_3$ ), the spectra difference more than twofold were labeled “+ ±-” (the boundary lines were  $y_2$  and  $y_4$ ). In this study, all proteins distributed outside the boundary line  $y_1$  and  $y_3$  were selected for subsequent analysis (Additional file 1: Fig S1C). Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) and Cluster of Orthologous Groups of proteins (COG) annotation based on <ftp://ftp.ncbi.nih.gov/gene/DATA/gene2go.gz>, <http://www.genome.jp/kegg> and <http://eggdb.embl.de/>, respectively. The GO and KEGG pathway enrichment analyses were performed using the R software.

## Interacting proteins co-expressed with prochymosin in *P. pastoris*

The DIPs PDI (Gene symbol: PAS\_chr1-1\_0160), BiP (Gene symbol: PAS\_chr2-1\_0140), Hsp70 (Gene symbol: PAS\_chr3\_0230), PEP4 (Gene symbol: PAS\_chr3\_1087) and PRB1 (Gene symbol: PAS\_chr1-1\_0226) were selected to co-expressed with prochymosin in *P. pastoris*. Fig. S2 depicts the plasmid vector structure carrying various interacting protein gene (co-expression vector). The restriction enzyme *EcoR* I is the location of insertion of interacting proteins. Vectors carrying different interacting protein genes were electrotransformed into strains chy-GS115 and chy34-GS115, respectively. Three positive clones were inoculated for each transform event and were cultured in 10 mL of BMGY with 0.1 mg/mL of zeocin and cultured at 28 °C and 200 rpm for 72 h. The cells were collected and resuspended in BMMY with 0.1 mg/mL of zeocin. The cultures were induced by 0.5% (v/v) methanol for 72 h. The supernatants were collected and used for comparing secretion levels of prochymosin by ELISA. The prime antibody and the second antibody were rabbit anti-prochymosin polyclonal antibody and goat anti-rabbit IgG conjugated to AP, respectively.

## Knockout of UGGT1 and UGGT2 genes by CRISPR/Cas9

Knockout of UGGT1 (GenBank: NC\_012965.1) and UGGT2 (GenBank: NC\_012963.1) genes of *P. pastoris* GS115 by CRISPR/Cas9 was according to [27]. The sgRNA was designed for the target site in the website E-CRISP (<http://www.e-crisp.org/E-CRISP/designcrisp.html>). The target sites were located at 800–819 bp of

UGGT1 gene and 344–363 bp of the UGGT2 gene. Gene knockout was verified by sequencing PCR products with genome as template, the sequencing primers were listed in Table 1.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12934-022-01904-3>.

**Additional file 1: Fig. S1** Polypeptide quality monitoring and differential interacting proteins screening data. **A** Peptide length distribution after trypsin digestion; **B** The number of spectra, peptides and proteins interacting with chy and chy34; **C** Distribution of differential interacting proteins (DIPs), red means the spectra difference more than 2-fold, yellow means the spectra difference more than 1.5-fold, grey means no difference. **Fig. S2** Map of the co-expression vector used to transformation of five DIPs into chy-GS115 or chy34-GS115 cell. pPFK indicates the promoter of 6-phosphofructo-1-kinase gamma-subunit from *P. pastoris*, AOX1 terminator indicates the alcohol oxidase transcriptional termination region, *EcoRI*s the integration site of DIPs. BleoR indicates the zeocin resistant gene, CEN6\_ARS indicates the yeast origin of replication, ori indicates the pUC origin. **Fig. S3** Verification of UGGT1 and UGGT2 knockout. A/C Sequence alignment, Sbjct and Query indicate the UGGT1/UGGT2 gene sequence before and after knockout. B/D Sequencing peak map of UGGT1/UGGT2 gene after knockout. The triangle indicate the deletion site. **Fig. S4** Growth curve acquisition of GS115, chy-GS115 and chy34-GS115 strains,  $\Delta$ UGGTs indicates that UGGT1 and UGGT2 have been knocked out. **Fig. S5 A** Far-UV CD spectra of wild and mutant prochymosins. The red line means the CD spectra of chy and the blue one was of the chy34. The table below showed the percentage of the protein's secondary structure of chy and chy34. **B** Fluorescence spectra of chy and chy34

**Additional file 2: Table S1.** Protein identification and abundance of differential interacting proteins with chy and chy34

**Additional file 3.** Classification of DIPs according to GO, KEGG and COG enrichment methods.

## Acknowledgements

Not applicable.

## Author contributions

NW and CY performed the experiments, analyzed the data, and wrote the manuscript. HP, WG, MW, GL performed the data curation, statistical analysis and article proofreading. DL and NW designed the study. All authors read and approved the final manuscript.

## Funding

This research was supported by National Natural Science Foundation of China [grant number: 31901661], Collaborative Innovation Project of Science and Technology of CAAS-GXSSA [CAAS-GXAAS-XTCX2019026-1] “Molecular design and breeding of important economic crops in Guangxi Province” and Foundation of Hainan Yazhou Bay Seed Lab [grant number: B21HJ0212].

## Availability of data and materials

All data generated or analysed during this study are included in this published article.

## Declarations

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no competing interests.

Received: 1 March 2022 Accepted: 21 August 2022  
Published online: 30 August 2022

## References

- Apweiler R, Hermjakob H, Sharon N. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim Biophys Acta*. 1999;1473(1):4–8.
- Shental-Bechor D, Levy Y. Effect of glycosylation on protein folding: a close look at thermodynamic stabilization. *PNAS*. 2008;105(24):8256–61.
- Adney WS, Jeoh T, Beckham GT, Chou YC, Baker JO, Michener W, Bruncky R, Himmel ME. Probing the role of N-linked glycans in the stability and activity of fungal cellobiohydrolases by mutational analysis. *Cellulose*. 2009;16:699–709.
- Price JL, Culyba EK, Chen W, Murray AN, Hanson SR, Wong CH, Powers ET, Kelly JW. N-glycosylation of enhanced aromatic sequons to increase glycoprotein stability. *Biopolymers*. 2012;98(3):195–211.
- Maldonado RF, Vieira DS, Alponi JS, Bonnell E, Thibault P, Ward RJ. Engineering the pattern of protein glycosylation modulates the thermostability of a GH11 xylanase. *J Biol Chem*. 2013;288(35):25522–34.
- Ge F, Zhu L, Aang A, Song P, Li W, Tao Y, Du G. Recent advances in enhanced enzyme activity, thermostability and secretion by N-glycosylation regulation in yeast. *Biotechnol Lett*. 2018;40(5):847–54.
- Han M, Yu X. Enhanced expression of heterologous proteins in yeast cells via the modification of N-glycosylation sites. *Bioengineered*. 2015;6(2):115–8.
- Zhang J, Wu J, Liu L, Li J. The crucial role of demannosylating asparagine-linked glycans in ERADicating misfolded glycoproteins in the endoplasmic reticulum. *Front Plant Sci*. 2021. <https://doi.org/10.3389/fpls.2020.625033>.
- Águila S, Noto R, Luengo-Gil G, Espín S, Bohdan N, de la Morena-Barrio ME, Peñas J, Rodenas MC, Vicente V, Corral J, Manno M, Martínez-Martínez I. N-Glycosylation as a tool to study antithrombin secretion, conformation, and function. *Int J Mol Sci*. 2021. <https://doi.org/10.3390/ijms22020516>.
- Helenius A, Aebi M. Roles of N-linked glycans in the endoplasmic reticulum. *Annu Rev Biochem*. 2004;73:1019–49.
- Caramelo JJ, Parodi AJ. A sweet code for glycoprotein folding. *FEBS Lett*. 2015;589(22):3379–87.
- Satoh T, Song C, Zhu T, Toshimori T, Murata K, Hayashi Y, Kamikubo H, Uchihashi T, Kato K. Visualisation of a flexible modular structure of the ER folding-sensor enzyme UGGT. *Sci Rep*. 2017. <https://doi.org/10.1038/s41598-017-12283-w>.
- Calles-García D, Yang M, Soya N, Melero R, Ménade M, Ito Y, Vargas J, Lukacs GL, Kollman JM, Kozlov G, Gehring K. Single-particle electron microscopy structure of UDP-glucose:glycoprotein glucosyltransferase suggests a selectivity mechanism for misfolded proteins. *J Biol Chem*. 2017;292(27):11499–507.
- Shin YJ, Vavra U, Veit C, Strasser R. The glycan-dependent ERAD machinery degrades topologically diverse misfolded proteins. *Plant J*. 2018;94(2):246–59.
- Dorner A, Kaufman R. The levels of endoplasmic reticulum proteins and ATP affect folding and secretion of selective proteins. *Biologicals*. 1994;22(2):103–12.
- Nishikawa SI, Fewell SW, Kato Y, Brodsky JL, Endo T. Molecular chaperones in the yeast endoplasmic reticulum maintain the solubility of proteins for retrotranslocation and degradation. *J Cell Biol*. 2001;153(5):1061–70.
- Shental-Bechor D, Levy Y. Effect of glycosylation on protein folding: a close look at thermodynamic stabilization. *Proc Natl Acad Sci USA*. 2008;105(24):8256–61.
- Li J, Yan Y, Wang Q, Sun L, Wang J. Effect of N-linked glycosylation on secretion and activity of recombinant DSPAa1 expressed in *Pichia pastoris*. *Chin J Biotechnol*. 2010;26(9):1287–92.
- Tian B, Chen Y, Ding S. A combined approach for improving alkaline acetyl xylan esterase production in *Pichia pastoris*, and effects of glycosylation on enzyme secretion, activity and stability. *Protein Expr Purif*. 2012;85:44–50.
- Han M, Wang X, Ding H, Jin M, Yu L, Wang J, Yu X. The role of N-glycosylation sites in the activity, stability, and expression of the recombinant elastase expressed by *Pichia pastoris*. *Enzyme Microb Technol*. 2014;54:32–7.
- Han M, Wang W, Jiang G, Wang X, Liu X, Cao H, Tao Y, Yu X. Enhanced expression of recombinant elastase in *Pichia pastoris* through addition of N-glycosylation sites to the propeptide. *Biotechnol Lett*. 2014;36(12):2467–71.
- Wang P, Zhang J, Sun Z, Chen Y, Liu JN. Glycosylation of prourokinase produced by *Pichia pastoris* impairs enzymatic activity but not secretion. *Protein Expr Purif*. 2000;20(2):179–85.
- Wang Z, Guo C, Liu L, Huang H. Effects of N-glycosylation on the biochemical properties of recombinant bEK L expressed in *Pichia pastoris*. *Enzyme Microb Technol*. 2018;114:40–7.
- Stefan RK, van den Brink Hans JM, Henrik RN, Zakaria F, Zdenko P, Egon BH, Eric J. Characterization of recombinant camel chymosin reveals superior properties for the coagulation of bovine and camel milk. *Biochem Biophys Res Commun*. 2006;342(2):647–54.
- Jensen JL, Mølgaard A, Poulsen JCN, Harboe MK, Simonsen JB, Lorentzen AM, Hjernø K, van den Brink JM, Qvist KB, Larsen S. Camel and bovine chymosin: the relationship between their structures and cheese-making properties. *Acta Crystallogr D Biol Crystallogr*. 2013;69:901–13.
- Wang N, Wang KY, Li Q, Guo WF, Liu DH. Expression and characterization of camel chymosin in *Pichia pastoris*. *Protein Expr Purif*. 2015;111:75–81.
- Generoso WC, Gottardi M, Oreb M, Boles E. Simplified CRISPR-Cas genome editing for *Saccharomyces cerevisiae*. *J Microbiol Methods*. 2016;127:203–5.
- Petrescu AJ, Milac AL, Petrescu SM, Dwek RA, Wormald MR. Statistical analysis of the protein environment of N-glycosylation sites: implications for occupancy, structure, and folding. *Glycobiology*. 2004;14(2):103–14.
- Yang M, Yu X, Zheng H, Sha C, Zhao C, Qian M, Xu Y. Role of N-linked glycosylation in the secretion and enzymatic properties of Rhizopus chinensis lipase expressed in *Pichia pastoris*. *Microb Cell Fact*. 2015. <https://doi.org/10.1186/s12934-015-0225-5>.
- Simone D, Masayuki I, Yutaka M, et al. Direct assay for endo- a -mannosidase substrate preference on correctly folded and misfolded model glycoproteins. *Carbohydr Res*. 2016;434:94–8.
- Conde R, Cueve R, Pablo G, Polaina J, Larriba G. A search for hyperglycosylation signals in yeast glycoproteins. *J Biol Chem*. 2004;279(42):43789–98.
- Young BP, Craven RA, Reid PJ, Willer M, Striling CJ. Sec63p and Kar2p are required for the translocation of SRP-dependent precursors into the yeast endoplasmic reticulum in vivo. *EMBO J*. 2001;20(1–2):262–71.
- Roversi P, Marti L, Caputo AT, Alonzi DS, Hill JC, Dent KC, Kumar A, Levasseur MD, Lia A, Waksman T, Basu S, Albrecht YS, Kristin Q, Mclvor JA, Lipp CB, Siliqi D, Vasiljević S, Mohammed S, Lukacik P, Walsh MA, Santino A, Zitzmann N. Interdomain conformational flexibility underpins the activity of UGGT, the eukaryotic glycoprotein secretion checkpoint. *Proc Natl Acad Sci USA*. 2017;114(32):8544–9.
- Takeda Y, Seko A, Fujikawa K, Izumi M, Kajihara Y, Ito Y. Effects of domain composition on catalytic activity of human UDP-glucose:glycoprotein glucosyltransferases. *Glycobiology*. 2016;26(9):999–1006.
- Modenutti CP, Capurro JIB, Ibba R, Alonzi DS, Song MN, Vasiljević S, Kumar A, Chandran AV, Tax G, Marti L, Hill JC, Lia A, Hensen M, Waksman T, Rushton J, Rubichi S, Santino A, Marti MA, Zitzmann N, Roversi P. Clamping, bending, and twisting inter-domain motions in the misfold-recognizing portion of UDP-glucose: glycoprotein glucosyltransferase. *Structure*. 2021;29(4):357–70.
- Fernández FS, Trombetta SE, Hellman U, Parodi AJ. Purification to homogeneity of UDP-glucose: glycoprotein glucosyltransferase from *Schizosaccharomyces pombe* and apparent absence of the enzyme from *Saccharomyces cerevisiae*. *J Biol Chem*. 1994;269(48):30701–6.
- Tripathi NK, Shrivastava A. Recent developments in bioprocessing of recombinant proteins: expression hosts and process development. *Front Bioeng Biotechnol*. 2019;7:420. <https://doi.org/10.3389/fbioe.2019.00420>.
- Lalonde ME, Durocher Y. Therapeutic glycoprotein production in mammalian cells. *J Biotechnol*. 2017;251:128–40.

39. Felberbaum RS. The baculovirus expression vector system: a commercial manufacturing platform for viral vaccines and gene therapy vectors. *Biotechnol J.* 2015;10(5):702–14.
40. Stephenne J. Development and production aspects of a recombinant yeast-derived hepatitis B vaccine. *Vaccine.* 1990;8:S69-73.
41. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta CT}$  method. *Methods.* 2001;25:402–8.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

