Quantitative N-glycoproteome analysis of bovine milk and yogurt

Jing Xiao, Jiqiu Wang, Renyou Gan, Di Wu, Yisha Xu, Lianxin Peng, Fang Geng

ABSTRACT
Post-translational modification structure of food’s proteins might be changed during processing, thereby affecting the nutritional characteristics of the food product. In this study, differences in protein N-glycosylation patterns between milk and yogurt were quantitatively compared based on glycopeptide enrichment, liquid chromatography separation, and tandem mass spectrometry analysis. A total of 181 N-glycosites were identified, among which 142 were quantified in milk and yogurt. Significant alterations in the abundance of 13 of these N-glycosites were evident after the fermentation of milk into yogurt. Overall, the N-glycosylation status of the majority of milk proteins remained relatively unchanged in yogurt, suggesting that their conformations, activities, and functions were maintained despite the fermentation process. Among the main milk proteins, N\textsuperscript{241} of cathepsin D and N\textsuperscript{558} of lactoperoxidase were markedly reduced after undergoing lactic acid fermentation to produce yogurt. Furthermore, a comparative analysis of current and previously reported N-glycoproteomic data revealed heterogeneity in the N-glycosylation of milk proteins. To sum up, a quantitative comparison of the N-glycoproteomes of milk and yogurt was presented here for the first time, providing evidence that the fermentation process of yogurt could cause changes in the N-glycosylation of certain milk proteins.

1. Introduction
Milk and dairy products are important food sources for humans, in particular newborns, because they contain essential nutrients, immune supplements, and bioactive ingredients (Sakandar and Zhang, 2021). In recent years, the extensive application of high-throughput proteomic technology in dairy research has greatly expanded our knowledge of milk proteins and provided novel information that has enhanced our understanding of milk’s properties. In this respect, one key development is that the post-translational modification of milk protein, especially N-glycosylation, has been the focus of many recent studies. As one of the most common types of post-translational modification, N-glycosylation covalently adds an oligosaccharide chain to the protein at an asparagine site. This oligosaccharide chain induces changes in the molecular weight, surface charge, and spatial conformation of the protein (Geng et al., 2017), which generally plays crucial roles in the folding and structural stability, activity maintenance, molecular recognition and interactions, and functional properties of the protein (Cambay et al., 2019). Moreover, N-glycosylation can also affect the digestion, nutritional properties, and health effects of food proteins. Therefore, the potential health risks or benefits of N-glycoproteins in milk constitute an important topic that deserves attention, yet they remain surprisingly understudied (Zheng et al., 2020). For example, N-glycosylation may weaken the hydrolysis of digestive enzymes through the steric hindrance of oligosaccharide chains, thereby prolonging the persistence of milk N-glycoproteins in the digestive tract, which could have non-negligible impacts on both intestinal immunity and intestinal microflora (Markussen et al., 2021). Accordingly, to explore and elucidate the mechanism by which milk N-glycoproteins affect human health, it is imperative to identify the N-glycosylation sites, oligosaccharide chain structures, and glycopeptide structures of milk N-glycoproteins to provide a robust foundation for evaluating their impact.

The differences in N-glycosylation of milk protein between differing mammal species and various milk components have been studied and compared. For example, Cao et al. (2019a) distinguished 58–100 N-glycoproteins in the whey of human/bovine colostrum and mature
N-glycoproteomes of bovine milk and yogurt were quantitatively elucidating the structures and functions of milk proteins. These findings suggest that protein N-glycosylation patterns exhibit pronounced variation among different sources and components of milk. During food processing, several factors e.g., heat, pH, ionic strength, and microorganisms, might induce deglycosylation, oxidation, and covalent cross-linking of oligosaccharide chains to participate in the Maillard reaction and significantly change the N-glycosylation modification of egg white proteins (Xiao et al., 2021). Yet, whether milk proteins’ N-glycosylation is altered during processing has not been investigated.

Yogurt is among the most popular dairy products because of its potential variation in texture and taste, improved digestibility and bioactivity, and lack of lactose (Jeffrey et al., 2020). During the fermentation process driven by lactic acid bacteria, milk proteins undergo acid-induced aggregation/denaturation in a low pH environment, leading to curdling (Zhao et al., 2021). Enzymes secreted by lactic acid bacteria trigger additional chemical alterations in milk proteins, such as their partial hydrolysis. Therefore, both chemical and biological factors may affect the frequency and extent of N-glycosylation of milk proteins during the processing of milk into yogurt products. In this study, the N-glycoproteomes of bovine milk and yogurt were quantitatively compared, aiming to uncover the alterations in N-glycosylation patterns of milk’s proteins during the production of yogurt. The results are expected to provide a new perspective and foundation for further revealing the nutritional properties and health benefits of yogurt from the proteins’ N-glycosylation.

2. Materials and methods

2.1. Preparation of yogurt

Yogurt starter (Lactobacillus bulgaricus and Streptococcus thermophilus, 1:1; Angel Yeast Co., Ltd, Yichang, China) was added to commercial bovine milk (Yili, ultra-high temperature treated milk, Tetra Pak packaging, 250 mL per package) at a ratio of 1:900 (w/v). After been magnetically stirred for 5 min, the starter was well mixed with milk, then the fermentation was conducted at 42 °C for 8 h, and fermented milk was incubated at 4 °C for 24 h. The yogurt was divided into tubes and stored at −80 °C until analysis. Bovine milk without a starter subjected to the same treatment was used as the control. Sampling was repeated three times to obtain biological replicates.

2.2. Protein extraction and enzymatic hydrolysis

Milk and yogurt samples were stored at −80 °C and total protein extraction was performed according to a previous study (Geng et al., 2018). Briefly, ultrasonic lysis was performed with the lysis buffer containing 8 mol/L urea and 1% protease inhibitor. Sample solutions were centrifuged at 4 °C and 12000 g for 10 min and the supernatant (total protein solution) was collected. Protein concentrations were determined by using the BCA protein assay kit (P0010, Beyotime Biotechnology Co., Ltd., Shanghai, China) (Geng et al., 2021). Proteins were reduced at 56 °C for 30 min with dithiothreitol (5 mmol/L) and alkylated in the dark at room temperature for 15 min with iodoacetamide (11 mmol/L). Finally, trypsin (V5280, Promega Corporation, Madison, WI, USA) was added at a mass ratio of 1:50 (trypsin: protein), and digestion was conducted overnight at 37 °C. Next, trypsin was added at a mass ratio of 1:100 (trypsin: protein), and the secondary hydrolysis was conducted for 4 h (Wang et al., 2022a).

2.3. LC-MS/MS analysis

Glycopeptides were enriched via hydrophilic interaction liquid chromatography and deglycosylated with peptide-N-glycosidase F (V4831, Promega Corporation, Madison, WI, USA) in H218O at 37 °C overnight (Xiao et al., 2020). Deglycopeptides were dissolved in solvent A (aqueous solution containing 0.1% formic acid and 2% acetonitrile) and separated using the LC system (EASY-nLC 1000, Thermo Fisher Scientific, Waltham, USA). The gradient was increased from 5% to 24% solvent B (aqueous solution containing 0.1% formic acid in 90% acetonitrile) for 40 min, 38% solvent B for 8 min, 80% solvent B for 4 min, and held at 80% solvent B for 4 min. The flow rate was maintained at 450 nL/min. Separated deglycopeptides were injected into the ion source for electrospray ionization (2.0 kV) and analyzed via tandem mass spectrometry (Orbitrap Fusion, Thermo Fisher Scientific) (Liu et al., 2020). Peptide precursor ions were fragmented via high-energy collision dissociation (35% of the fragmentation energy) and data were obtained using the data-dependent acquisition mode (Yang et al., 2020). Other mass spectrometry parameters were the same as previously reported (Wang et al., 2021a).

2.4. Protein identification and quantitative analysis

Mass spectral data were processed using MaxQuant (v1.5.2.8). UniProt Bos taurus (bovine) (37,948 sequences) databases were used to retrieve sequences. Specifically, mass tolerance was set as 5 ppm for precursor ions in the main search and 0.02 Da for fragment ions (Xin et al., 2020). Cysteine alkylation was set as fixed modification and variable modifications were methionine oxidation, protein N-terminal acetylation, deamidation (NQ), and asparagine deamidation (Nφ) (Liu et al., 2021). The false discovery rate for protein and peptide identification was set to 1%.

The intensity of modified peptides in each sample was determined using the label-free quantitative calculation method and the relative quantitative value (ratio) of each sample was obtained according to the intensities of modified peptides between different groups (Luo et al., 2021). At p < 0.05, fold change (FC) exceeding 1.3 was used as the threshold for the significant increase and <0.77 as the threshold for the significant decrease.

2.5. Identification of proteins from lactic acid bacteria

Another search was separately performed by using the same MS/MS raw data to identify the proteins from lactic acid bacteria in the yogurt sample. Briefly, UniProt databases were used to retrieve sequences from Lactobacillus bulgaricus and Streptococcus thermophilus (36737 sequences). Mass tolerance was set as 5 ppm for precursor ions in the main search and 0.02 Da for fragment ions, and the false discovery rate for protein and peptide identification was set to 1%.

2.6. Bioinformatics and multivariate analysis

Bovine milk and yogurt N-glycoproteins were annotated using the UniProt-GOA database (http://www.ebi.ac.uk/GOA/) and InterProScan (v.5.14–53.0, http://www.ebi.ac.uk/interpro/) in the gene ontology analysis. First, protein ID was converted to UniProt ID, next UniProt ID was used to match Gene Ontology (GO) ID, and the corresponding information was retrieved from the UniProt-GOA database according to
the repeated peptide sequences, 172 unique sequences with N-glyco-
MS/MS analysis, of which 8531 spectra were matched to 742 peptide
3.1. Overview of the identified N-glycoproteome
sequences, accessed 2019.09.15). The mass error of all peptides was <5
ppm, indicating our test results were accurate (Fig. 1A). After removing
the repeated peptide sequences, 172 unique sequences with N-glyco-
sylation were obtained. These N-glycopeptides belonged to 118 N-gly-
coproteins and contained 181 N-glycosites (Table S1). Among the
identified N-glycoproteins, 81 (68.6%) were single N-glycosylated, 25
(21.2%) were double N-glycosylated, and 12 (10.2%) were multiple N-
glycosylation sites per identified N-glycoprotein.

A total of 144143 secondary mass spectra were obtained in the LC-
MS/MS analysis, of which 8531 spectra were matched to 742 peptide
sequences by searching the cattle database (Bos taurus, 37,948 se-
quencies, accessed 2019.09.15). The mass error of all peptides was <5
ppm, indicating our test results were accurate (Fig. 1A). After removing
the repeated peptide sequences, 172 unique sequences with N-glyco-
sylation were obtained. These N-glycopeptides belonged to 118 N-gly-
coproteins and contained 181 N-glycosites (Table S1). Among the
identified N-glycoproteins, 81 (68.6%) were single N-glycosylated, 25
(21.2%) were double N-glycosylated, and 12 (10.2%) were multiple N-
glycosylation sites per identified N-glycoprotein; NXS, whereas the proportion of other non-conserved motifs was
3. Results and discussion
3.2. Functional annotation and enrichment analyses of identified N-
glycoproteins
GO annotations were conducted to evaluate the functions of the
identified N-glycoproteins. Under the “biological process” category, N-
glycoproteins were mainly involved in “single-organism process”
(GO:0044700, 44 N-glycoproteins), “biological regulation”
(GO:0065007, 38 N-glycoproteins), “cellular process” (GO:0009987, 37
N-glycoproteins), “metabolic process” (GO:0008152, 33 N-glycopro-
tins), and “response to stimulus” (GO:0050896, 27 N-glycoproteins).
Under the “cellular components” category, “extracellular region”
(GO:0005576, 47 N-glycoproteins), “cell” (GO:0005623, 44 N-glyco-
proteins), “organelle” (GO:0043226, 44 N-glycoproteins), and “mem-
brane” (GO:0016020, 37 N-glycoproteins) were the GO terms
predominantly associated with most N-glycoproteins. Further, the
majority of identified N-glycoproteins were annotated with functions of
“binding” (GO:0005488, 64 N-glycoproteins) and “catalytic activity”
(GO:0003824, 29 N-glycoproteins) (Fig. 2A). Next, the putative sub-
cellular localization of every identified N-glycoprotein was predicted by
using Wolfsort (http://wolfpsort.seq.cbrc.jp/). Subcellular localization
annotations were obtained for 96 N-glycoproteins, half of which featured “extracellular” localization. This finding is consistent with the
secretion of bovine milk by mammary gland cells. Although 19 N-gly-
coproteins were located to the “plasma membrane”, relatively few were
found present in other subcellular structures (Fig. 2B).

We also conducted GO enrichment analysis of the identified N-gly-
coproteins to further explore their functions and associated processes.
Consistent with their GO classifications, the “extracellular region” was
overrepresented at GO level 2 with a fold-enrichment value of 3.66. The
“extracellular space,” “cell surface,” “vesicle,” “extracellular organelle,”
and “endomembrane system” terms were enriched at GO level 4
(Table S2). Under the “molecular function” category, “glycosaminogly-
can binding” (GO:0005539) was enriched with a fold-enrichment
value of 12.46 and mapped to six N-glycoproteins. Two enzymes

Fig. 1. Characteristics of identified N-glycoproteins from bovine milk and yogurt. (A) Mass error distribution of the identified N-glycopeptides; (B) number of N-
glycosylation sites per identified N-glycoprotein; (C) the sequence motifs of the identified N-glycopeptides; (D) distribution of recognized sequence motifs based on
the unique sequences or MS/MS counts of the identified N-glycopeptides (X ≠ P).
MFGM (Cao et al., 2019b). The term (GO:0009617, level 6), with 7 mapped N-glycoproteins, was overrepresented (immunoglobulin J chain, Q3SYR8; monocyte differentiation antigen CD14, Q95122; toll-like receptor 2, Q95LA9), antimicrobial substance represented with a fold-enrichment value of 7.69. These N-glycoproteins are involved in signal transduction and regulation during the response to bacterium, mainly via the immune system (GO:0006952, level 6), 13 mapped N-glycoproteins. Many N-glycoproteins were annotated with the GO term (A0A3Q1MDS4, N-glycosite), glycoprotein 2 (A0A3Q1M193, N-glycosite), procollagen-lysine, -proline residues in milk protein as substrates. Hence, it is theoretically possible to produce hydroxyproline by itself in milk, which could lead to uncertainty in the identification of adulterated milk (Liu et al., 2014). The two enzymes identified in this study catalyze the production of hydroxyproline using the abundant proline residues in milk protein as substrates. However, it is very low in fresh milk, relatively high levels have been detected in adulterated milk (Liu et al., 2014). The two enzymes identified in this study catalyze the production of hydroxyproline using the abundant proline residues in milk protein as substrates. Hence, it is theoretically possible to produce hydroxyproline by itself in milk, which could lead to uncertainty in the identification of adulterated milk. "Endopeptidase inhibitor activity" (GO:0004686, GO level 6) was also overrepresented, with a fold-enrichment value of 7.2 and mapped to five N-glycoproteins. These inhibitors may suppress the hydrolysis of milk proteins and thus positively affect the stability and storage of milk.

Under the category “biological process,” “cell surface receptor signaling pathway” (GO:0007166, level 6) was overrepresented with a fold-enrichment value of 2.99 and 13 mapped N-glycoproteins. Many N-glycoproteins are involved in signal transduction and regulation during lactic acid bacteria fermentation, some of which enter milk in the form of milk exosomes or MFGM (Cao et al., 2019b). The term “response to bacterium” (GO:0009617, level 6), with 7 mapped N-glycoproteins, was overrepresented with a fold-enrichment value of 7.69. These N-glycoproteins defend against bacterial invasion, mainly via the immune system (immunoglobulin J chain, Q3SYR8; monocyte differentiation antigen CD14, Q95122; toll-like receptor 2, Q95L Aj9), antimicrobial substance production (lactoperoxidase, P80025), and iron-binding (lactotransferrin, P24627). Notably, the transforming growth factor beta-1 proprotein (TGFβ1, P18341) was associated with several enriched GO terms, including “regulation of bone mineralization” (GO:0045778 and GO:0030500), “regulation of collagen metabolic and biosynthetic processes” (GO:0010714 and GO:0032965), “regulation of fibroblast proliferation” (GO:0048145), and “response to progesterone” (GO:0032570). TGFβ1, an important cytokine, regulates the growth and differentiation of various cell types and participates in multiple physiological processes. For example, TGFβ1 activates the cAMP-responsive element-binding protein 3-like protein 1 by regulating the intracellular proteolysis to stimulate the continuous production of collagen, and is also an effective stimulator of osteoblast bone formation, leading to chemotaxis and the proliferation and differentiation of definitive osteoblasts (Ye, 2020). Although TGFβ1 is expressed by many cell types, the activated protein (i.e., fragment 279–390) exerts a localized effect in the cellular environment due to regulation by the latency-associated peptide chain (i.e., fragment 30–276) and “environmental molecules” (Dubois et al., 1995). In our study, two N-glycoproteins of milk TGFβ1 were identified (N82 and N136), both being located in the latency-associated peptide chain fragment; this evidence supports the involvement of N-glycosylation in the activation of TGFβ1.

### 3.3. Quantitative comparison of N-glycoproteomes between milk and yogurt

The N-glycoproteomic profiles of milk and yogurt samples were initially compared using PCA (principal component analysis). The three replicate samples in each milk or yogurt group clustered together, indicating good reproducibility of the quantitative analysis, whereas the distribution of the two groups was distinct (Fig. 3A), indicating marked differences in the N-glycoproteome between milk and yogurt. Based on the signal intensities of the identified N-glycopeptides, the abundances of N-glycosites and N-glycoproteins were quantified. Of the 181 N-glycosites and 118 N-glycoproteins we identified, 142 N-glycosites on 94 N-glycoproteins were quantified (Fig. 3B). Using the milk group as a reference, N-glycosites and N-glycoproteins with p < 0.05 and fold-change > 1.3 or < 0.77 were respectively classified as significantly upregulated and downregulated. Overall, 13 N-glycoproteins (containing 13 N-glycosites) in yogurt were significantly altered in comparison to milk (Table S3). Five N-glycoproteins (containing 5 N-glycosites) were markedly upregulated, namely Ras-related protein Rab-3 (A0A3Q1MD54, N-glycosite), glycoprotein 2 (A0A3Q1M193, N-glycosite), S-nucleotidase (Q5927, N-glycosite), lipocalin 2 (E1B626, N-glycosite), procollagen-lysine, -proline, and -lysino dipeptide (A0A3Q11S80, N-glycosite) (Fig. 4A), while eight N-glycoproteins (containing 8 N-glycosites) were significantly downregulated (Fig. 4B). Among these, lactoperoxidase (P80025, N-glycosite, fold-change: 0.46) and cathepsin D (P80209, N-glycosite, fold-change: 0.47) were the most significantly downregulated.

During the fermentation process to produce yogurt, many factors might affect the N-glycosylation of milk proteins. The physiological and metabolic activities of lactic acid bacteria would be the potentially major reasons for the changed N-glycosylation of milk proteins. Although lactic acid bacteria are incapable of protein post-translational processing, there is evidence they could secrete proteases to hydrolyze milk proteins (Gu et al., 2020). To assess the influence of lactic acid bacteria upon milk protein, proteins from L. bulgaricus and S. thermophilus were mined from M5/MS data. Through these results, we identified a total of 39 proteins from L. bulgaricus and S. thermophilus (Table S4). Among them, two kinds of proteolytic enzymes came from S. thermophilus: peptidase (A0A2X3UL7) and serine protease (V8B7RF). A glycosyltransferase (U4M5H4) from S. thermophilus was also identified, which implied that the ends of oligosaccharide chains on milk proteins might be processed by lactic acid bacteria. Admittedly, the number of identified lactic acid bacteria proteins in this study was very limited because their abundance in the yogurt samples was generally very low. Nevertheless, the limited information that can be gleaned from
the present study indicates that lactic acid bacteria are likely capable of changing the N-glycosylation profile of protein in milk during its fermentation to make yogurt.

The Maillard reaction during milk fermentation could also change the N-glycosylation of milk protein. This reaction is known to change the N-glycosylation of egg white protein, in that the abundance of 18 N-glycoproteins (containing 32 N-glycosites) was significantly reduced in the spray-dried egg white compared with fresh egg white; hence, the N-glycans of egg white proteins were likely covalently cross-linked during spray drying (Liu et al., 2021). Unlike spray drying, the process of fermentation for yogurt is mild. Nevertheless, the Maillard reaction may ensue under very mild conditions, even under physiological ones, such as the glycation of plasma proteins (Rossello et al., 2021). It’s conducive to the Maillard reaction’s initiation in the early stage of fermentation for yogurt: high reducing sugar content, abundant protein content, and a temperature of 40°C–42°C. Therefore, whether the oligosaccharide chain on milk protein is changed due to the Maillard reaction during the fermentation for yogurt is another question that needs to be answered in

Fig. 3. Quantitative comparison of protein N-glycosylation between the bovine milk and yogurt samples. (A) Principal component analysis of milk (M) and yogurt (Y) N-glycoproteomes; (B) the number of differentially abundant N-glycoproteins and N-glycosites between milk and yogurt.

Fig. 4. Variation in the abundances of N-glycoproteins (N-glycosites) between bovine milk (M) and yogurt (Y); Results are showing by mean ± SE, n = 3; *p < 0.05, **p < 0.01.
future investigations. To sum up, our results confirmed that the N-glycosylation of milk protein was changed during yogurt’s fermentation process. The underlying mechanism was speculated upon but awaits verification in subsequent studies.

3.4. Changes in N-glycosylation of main milk proteins during yogurt’s fermentation

3.4.1. Lactotransferrin

Lactotransferrin (P24627) was the most abundant N-glycoprotein identified in the present study. A search of the UniProt database revealed the presence of four marked N-glycosites (N^106, N^309, N^59, and N^649). We succeeded in identifying the N^252, N^212, and N^564 sites other than the N^387. An earlier study by Hao et al. (2019a) revealed different N-glycosites of lactotransferrin in bovine colostrum milk (N^252, N^253, N^564, N^572, and N^649) and bovine mature milk (N^252, N^253, N^649, and N^782); this implied the N-glycosylation of lactotransferrin is affected by the stage of lactation and processing method. During the processing of bovine milk into yogurt, N-glycosylation of lactotransferrin was slightly increased but the change was not significant (FC = 1.04–1.22, p = 0.11–0.74; Table S1). The results suggested that the fermentation process affects the N-glycosylation of lactotransferrin.

Lactotransferrin is a monomeric glycoprotein, with a molecular weight of 78 kDa, which binds two Fe^3+ and two CO_3^2- (Moore et al., 1997) and has multiple biological functions. These include regulating iron absorption and immune responses (Hao et al., 2019) as well as engaging in antibacterial, antiviral, anti-inflammatory, and anti-cancer activities (Zhang et al., 2015b). It is currently believed that the antibacterial activity of lactotransferrin is mainly determined by two factors. On the one hand, lactotransferrin can inhibit the growth of bacteria thus has an antibacterial effect, which is related to its ability to chelate free iron; on the other hand, lactotransferrin directly induces the release of lipopolysaccharide in the outer membrane of bacteria, thereby also inhibiting bacterial growth. Further, lactotransferrin exerts assimilation, differentiation, and anti-apoptotic effects on osteoblasts, inhibits osteoclast production, and regulates bone growth (Shi et al., 2020). The covalently modified N-glycans on lactotransferrin will influence its binding to potential interacting molecules, such as receptors, thus playing a pivotal role in determining its activities.

3.4.2. Lactoperoxidase

We identified a total of five N-glycosites (N^106, N^212, N^252, N^358, and N^499) of lactoperoxidase (P80025). In the UniProt database, all five sites are annotated. The N^106 of lactoperoxidase is predicted to be N-glycosylated, which was experimentally validated in this study, and the four other sites were reported on previously (Singh et al., 2009). In the work by Hao et al. (2019a), two N-glycosites of lactoperoxidase were identified in bovine mature milk (N^12 and N^564) but only one (N^564) in bovine colostrum; this suggests the N-glycosylation of lactoperoxidase is heterogeneous and varies with lactation and processing. Quantitative comparisons in our study uncovered changes to the magnitude of N-glycosylation of lactoperoxidase after fermentation into yogurt. While the abundances of N^358 (FC = 0.46, p < 0.01), N^12 (FC = 0.63, p = 0.38), and N^499 (FC = 0.87, p = 0.20) were lower in yogurt than milk, only the reduction in N^358 was significant. The N-glycosite of N^106 remained unaffected (FC = 1.02, p = 0.80). The N^252 of lactoperoxidase was identified in yogurt only, whereas the other four sites were detected in both milk and yoghurt samples. Our data provide compelling support for altered N-glycosylation patterns of lactoperoxidase during yogurt’s fermentation process that could impact protein activity and function.

Lactoperoxidase is an enzyme naturally present in milk, with a molecular weight of ~80 kDa, that functions in the production of the antimicrobial compound hypothiocyanous acid (HOSCN) in the presence of thiocyanate (SCN) and H_2O_2 (Jafary et al., 2013). Therefore, lactoperoxidase is considered the major antibacterial ingredient that protects the udder from infection and delays the degradation of milk. For example, lactoperoxidase can extend the shelf life of raw milk to more than 16 h at 40 °C (Masud et al., 2008). In addition, lactoperoxidase could significantly reduce IL-8 mRNA expression and inhibit secretion of IL-8 induced by H_2O_2 in Caco-2 cells, leading to reduced intestinal inflammation associated with oxidative stress (Matsushita et al., 2008).

Developing effective medications using lactoperoxidase as an ingredient have been reported and its stable nanoformulations—prepared by loading or adsorbing bovine lactoperoxidase onto chitosan nanoparticles—display selective apoptosis-mediated anticancer activity (Abu-Serie and El-Fakhrany, 2017). However, the potential relevance of post-translational N-glycosylation for lactoperoxidase activity has yet to be established. The N-glycosylation patterns of lactoperoxidase in milk are diverse among species, not only in terms of N-glycosites but also N-glycan profiles (Kumar et al., 2018). These differences in N-glycosylation structures may be a key underlying reason for the differences in both the activity and characteristics of lactoperoxidase among various mammalian species. Accordingly, clarifying the relationship between N-glycosylation and lactoperoxidase function is now an important research goal.

3.4.3. Mucin-15

As the main component of epidermal mucus, mucins are proteins that usually have a large molecular weight and a high degree of glycosylation and provide a physicochemical barrier to protect epithelial tissues. Several mucins have been identified in milk, of which mucin-15 is one of the major MFGM proteins that accounts for 0.08% of the milk’s protein content and 1.5% of all its MFGM proteins (Smoczyn’zki et al., 2011). Western blot analysis of mucin-15’s distribution in different components of bovine milk revealed its presence in both fat-containing components, such as MFGM and non-fat components, such as skimmed milk and whey. The glycans of mucin-15 consist of fucose, galactose, mannose, sialic acid, and both N-acetylgalactosamine and N-acetylmuramylglycamin (Pallesen et al., 2007).

In this study, six N-glycosites (N^30, N^44, N^54, N^71, N^79, and N^221) were identified in mucin-15 (Q8M011). Quantitative data showed that after fermentation into yogurt, the abundances of four of these N-glycosites of mucin-15 were increased (N^30, FC = 1.44; N^44, FC = 1.13; N^71, FC = 1.11; N^79, FC = 1.11) but decreased for the other two sites (N^54, FC = 0.98; N^271, FC = 0.93). However, these changes were not significant (p = 0.87–0.10). These results indicate an overall upward trend for the N-glycosylation of mucin-15 (average FC = 1.12) during milk’s fermentation into yogurt. These changes to mucin-15’s N-glycosylation may be attributed to the action of related enzymes in milk or starter cultures (L. bulgaricus and S. thermophilus). Highly N-glycosylated mucin-15 is expected to produce several glycopeptides during the gastrointestinal digestion process and further studies are essential to establish the effects of these glycopeptides on intestinal immunity and flora.

3.4.4. Cathepsin D

Cathepsin D (P80209), an acid endoprotease, was annotated using the UniProt database as a protein whose catalytic specificity is similar to that of pepsin A. Annotation of GO cellular components disclosed the presence of bovine milk cathepsin D in “lysosome” (GO: 0005764), indicating its derivation from lysosomes of breast cells and subsequent transport into milk during lactation. Human cathepsin D mRNA has been detected in milk products. Furthermore, a number of human milk peptides identified via mass spectrometry are predicted to undergo cleavage by cathepsin D (Khalidi et al., 2014). Therefore, cathepsin D is considered the main protease catalyzing proteolysis in human milk. Based on the high homology between bovine and human cathepsin D (79.4% shared identity determined via the Align analysis; https://www.uniprot.org/align/), bovine cathepsin D was similarly inferred to be one of the main enzymes contributing to milk proteolysis.

Bovine cathepsin D is annotated with two potential N-glycosites. In the current study, however, only one N-glycosite (N^241) of cathepsin D (P80209) was identified. After fermentation into yogurt, the abundance
of the cathepsin D N\textsuperscript{241} site was significantly decreased (p < 0.01) with an FC of 0.47. This reduction in the degree of N-glycosylation has a significant potential impact on the activity of bovine cathepsin D in yogurt, although its active hydrolysis sites (N\textsuperscript{77} and N\textsuperscript{273}) are distant from the N-glycosite (N\textsuperscript{241}). In the crystal structure of human cathepsin D, an N-glycan extends from Asn-70 towards Lys-203 and is potentially involved in the recognition regions of phosphotransferase and human cathepsin D (Metcalf and Fuesk, 1993). Furthermore, N-glycans may affect both the thermal stability and water activity tolerance of proteases. The potential role of N\textsuperscript{241} glycosylation in bovine cathepsin D enzyme activity now requires further exploration.

3.4.5. Folate receptor α

Folate in bovine milk coexists with folate-binding protein (mainly folate receptor α), in the form of a complex. The folate-binding affinity of human folate receptor α is not significantly affected by enzymatic deglycosylation; however, when the folate receptor α is deglycosylated via mutation it shows a decrease or complete loss of its folate-binding capacity (Roberts et al., 1998). These results indicate that the N-glycans of folate receptor α are not necessary for folate binding but essential for synthesis in the correct conformation. Previous studies have shown that bovine folate receptor α (P02702) from milk whey (Cao et al., 2019a) and MFGM (Cao et al., 2019b) contains three N-glycosites (N\textsuperscript{62}, N\textsuperscript{58}, and N\textsuperscript{160}), all of which were successfully identified in our study. Notably, the N-glycosylation levels of N\textsuperscript{160} (FC = 0.76, p = 0.018) were significantly decreased after yogurt’s fermentation process. Still, changes in N-glycosylation may have little effect on the binding of folate receptor α to folate given the homology between human and bovine folate receptor α (Roberts et al., 1998). The acid produced by fermentation may induce dissociation of the complex, so that folate in yogurt mainly exists in a free state (Laiò et al., 2013). In any case, yogurt’s fermentation process is gentle, and therefore the folate content is maintained and even increased if the starter bacterium used is a probiotic that can synthesize folate (Wang et al., 2021b). Compared with fermentation for yogurt, thermal processing of milk is disadvantageous for the retention of folate in milk. An earlier study has shown that the complex of folate and folate receptor α in milk is impaired by thermal processing. Specifically, the complex is stable in pasteurized milk but folate will dissociate and undergo thermal degradation during UHT milk processing. Heating at temperatures below 80 °C, however, is safe for ensuring the complex retains its integrity (Sahoo et al., 2014).

3.4.6. Cluster of differentiation (CD) antigens

A total of seven CD antigens were identified as N-glycosylated in this study: platelet glycoprotein 4 (CD36, P26201; N\textsuperscript{77}, N\textsuperscript{172}, N\textsuperscript{235}, N\textsuperscript{247}, N\textsuperscript{321}, N\textsuperscript{417}), monocyte differentiation antigen CD14 (CD14, Q95122; N\textsuperscript{150}, N\textsuperscript{280}), CD44 antigen (Q29423, N\textsuperscript{97}), integrin-associated protein (CD47, F1M1U0; N\textsuperscript{40}, N\textsuperscript{60}, N\textsuperscript{99}), CD63 antigen (Q9XS2K; N\textsuperscript{330} and N\textsuperscript{358}), tetraspanin (CD82, A0A3Q1NBQ9; N\textsuperscript{197}), and CD109 (A0A3Q1M299; N\textsuperscript{11}, N\textsuperscript{114}, and N\textsuperscript{353}). Platelet glycoprotein 4 is one of the major immune-related N-glycoproteins in bovine milk. During the processing of yogurt, glycosylation of N\textsuperscript{172} of platelet glycoprotein 4 was increased, but not to a significant extent (FC = 1.45, p = 0.10), while the other five N-glycosites showed a decreasing trend, which was also not significant (FC = 0.78–0.97; p = 0.41–0.82). Being a coreceptor for Toll-like receptor 4, Toll-like receptor 6, and the Toll-like receptor 6 heterodimer, platelet glycoprotein 4 in combination oxidized low-density lipoprotein or micrornas diacylated lipoprotein to form CD36 clusters and initiate signal transduction. These activated signaling pathways are known to be involved in the inflammatory response. The platelet glycoprotein 4 content in milk with a higher number of somatic cell counts (<4 × 10\textsuperscript{5} cells/mL, closely related to mastitis) was significantly altered when compared to milk with a low somatic cell count (<10\textsuperscript{5} cells/mL, normal milk) (Zhang et al., 2015a). In addition, platelet glycoprotein 4 was annotated as a participant in the absorption and metabolism of lipids, such as triglyceride transport (GO:0034197), intestinal cholesterol absorption (GO:0030299), long-chain fatty acid import into cells (GO:0044539), and cholesterol import (GO:0070508). In view of these potential functions, platelet glycoprotein 4 is proposed to figure prominently in the process of lipid absorption of milk during its digestion in the small intestine.

CD14 is a coreceptor of lipopolysaccharides that together with lipopolysaccharide-binding protein binds to lipopolysaccharide and activates Toll-like receptor 4 and its downstream immune response. Therefore, CD14 could act as an immune effector and participate in the defense against invasion by microorganisms. The abundance of CD14 in milk is significantly reduced during thermal processing, whereas ultrasonic treatment and microfiltration treatment may be effectively used to maintain CD14 levels in milk (Zhang et al., 2021). Changes in the abundance of immune and antibacterial proteins in milk caused by different processing methods may be related to the shelf life of milk.

CD44 antigen is a hyaluronic acid-binding receptor that can be activated or upregulated in immune cells through antigen receptor stimulation, thereby inducing proinflammatory cytokine production. In addition to participating in immune and inflammation-related regulation, CD44 antigen plays a key regulatory role in the lipid metabolism of bovine mammary epithelial cells. Both triglyceride and cholesterol contents are augmented in cells with overexpression and conversely reduced in cells with knockdown of CD44 (Jiang et al., 2020), thus suggesting its critical role in regulating lipid components of bovine milk. The hyaluronic acid-binding domain of CD44 antigen is N-glycosylated and the N-glycans of this domain could weaken hyaluronic acid binding upon capping the terminal with sialic acid residues (Guvench, 2015). This weakened binding is due to interactions of the terminal sialic acid of N-glycan with the basic residue on the hyaluronic acid-binding domain, to occupy the binding site. Our results indicate a critical role of covalently bound N-glycans in the binding of CD44 antigen to its ligand.

After fermentation into yogurt, the abundances of all 18 N-glycosites of the above seven CD antigens did not significantly alter, indicating that the N-glycosylation level of these immune-related proteins remained relatively stable during the yogurt fermentation process. The results suggest that these proteins remain active and potentially contribute to the health benefits of yogurt.

3.5. Heterogeneity of the milk N-glycoproteome

Heterogeneity is one of the important characteristics of glycosylation. Our research group previously demonstrated the tissue specificity of N-glycosylation (Geng et al., 2017; Xiao et al., 2020; Yang et al., 2020), though other research groups have reported differing N-glycosites of the same protein. Additionally, heterogeneity in the N-glycosylation of milk protein has been documented. We compared our results with the N-glycoproteomes of MFGM and whey (Cao et al., 2019a; Cao et al., 2019b). Overall, MFGM contained the most types of N-glycoproteins and N-glycans (176 N-glycoproteins having 273 N-glycosites), followed by the milk samples of this study (118 N-glycoproteins having 181 N-glycosites), and milk whey contained the least (98 N-glycoproteins having 139 N-glycosites). The overall differences in N-glycoproteomes among the three samples could be mainly attributed to their variable degrees of homogenization of protein. The presence of high-abundance proteins usually reduces the throughput of proteome identification, which is relatively low in MFGM.

Besides the overall differences in N-glycoproteomes between samples, the number and location of N-glycosites of the same N-glycoprotein varied among the samples as well. Platelet glycoprotein 4 (P26201), glycoprotein 2 (A0A3Q1M193), and mucin-15 (Q8M101) are used here as examples. The number of N-glycosites of platelet glycoprotein 4 from whey, milk, and MFGM respectively were 7, 2, and 6. Among these sites, N\textsuperscript{172} and N\textsuperscript{321} were each shared by three types of platelet glycoprotein 4 but N\textsuperscript{305} was unique to platelet glycoprotein 4 from MFGM (Fig. 5A).
Similarly, glycoprotein 2 from whey had only one identified N-glycosite (N\textsubscript{257}) while that from milk harbored 6 N-glycosites. Three additional N-glycosites (N\textsubscript{33}, N\textsubscript{142}, and N\textsubscript{339}) were identified on glycoprotein 2 from MFGM (Fig. 5B). Taken together, these results are consistent with our data for chicken egg N-glycoproteomes showing different N-glycosite numbers (15–28) and locations of chicken mucin-5B from differing egg parts (Xiao et al., 2020). This is an “economical and efficient” mode acquired by living organisms in the long-term evolutionary process: a protein template, yielding multiple structures, with various activities.

4. Conclusions

A total of 181 N-glycosites on 118 N-glycoproteins were identified in bovine milk and fermented yogurt. The quantification of the abundance of 142 N-glycosites revealed significant changes in 13 of them (5 upregulated, 8 downregulated) after fermenting milk into yogurt. Overall, the N-glycosylation profile of milk proteins was only slightly changed during the process of lactic acid fermentation to produce yogurt. For example, N-glycosylation patterns of the main milk proteins, lactotransferrin, and mucin-15 showed no significant differences. These findings suggest that a number of N-glycosylation sites did undergo significant changes in abundance, which warrants closer attention. For instance, N\textsubscript{241} of cathepsin D and N\textsubscript{358} of lactoperoxidase were significantly downregulated during the yogurt fermentation process, suggesting the reduction of N-glycans at the molecular surface could trigger changes in their surface charges and properties. Our quantitative analysis of the protein N-glycosylation patterns between milk and yogurt provide novel information that should facilitate a better understanding of the molecular structural changes of proteins during the fermentation of milk into yogurt.

CRediT authorship contribution statement

Jing Xiao: Investigation, Data curation, Writing – original draft.
Jinqiu Wang: Data curation, Writing – original draft, Writing – review & editing.
Renyou Gan: Resources, Writing – review & editing.
Di Wu: Methodology, Data curation.
Yisha Xu: Methodology, Data curation.
Lianxin Peng: Methodology, Resources, Funding acquisition.
Fang Geng: Writing – original draft, Writing – review & editing, Supervision.

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.

Acknowledgments

This work was supported by the National Key Research and Development Program of China (2018YFC1602101 and 2018YFC1604301).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.
Cambay, F., Henry, O., Durocher, Y., De Crescenzo, G., 2019. Impact of N-glycosylation on IgG receptor/IGG interactions: unravelling differences with an enhanced surface plasmon resonance biosensor assay based on coiled-coil interactions. mAbs 11 (3), 435–452.

Cao, X.Y., Yang, M., Yang, N., Liang, X.N., Tan, D.B., Liu, B., Wu, J.R., Yue, X.Q., 2019a. Characterization and comparison of whey N-glycopeptides from human and bovine colostrum and mature milk. Food Chem. 276, 266–273.

Cao, X.Y., Zheng, Y., Wu, S.Y., Yang, N., Wu, J.R., Liu, B., Ye, W.H., Yang, M., Yue, X.Q., 2019b. Characterization and comparison of milk fat globule membrane N-glycopeptides from human and bovine colostrum and mature milk. Food Funct. 10 (8), 5046–5058.

Dobius, C.M., Laprise, M.-H., Blanchette, F., Gentry, L.E., Leduc, R., 1995. Processing of transforming growth factor β1 precursor by human furin convertase. J. Biol. Chem. 270 (18), 10618–10624.

Geng, F., Wang, J.Q., Liu, D.Y., Jin, Y.G., Ma, M.H., 2017. Identification of N-glycotypes in chicken egg yolk proteins using an omics strategy. J. Agric. Food Chem. 65 (26), 5357–5364.

Geng, F., Xie, Y., Wang, W., Wang, J., 2021. Depolymerization of chicken egg yolk granules induced by high-intensity ultrasound. Food Chem. 354, 129580.

Gu, Y., Li, X., Liu, H., Li, Q., Xiao, R., Duda, O.E., Yang, L., Ma, Y., 2020. The impact of multiple species starters on the peptide profiles of yoghurts. Int. Dairy J. 106, 104684.

Gowrishankar, D., 2015. Revealing the mechanisms of protein disorder and N-glycosylation in CD44 hyaluronan binding using molecular simulation. Front. Immunol. 6, 305.

Hao, L.Y., Shan, Q., Wei, J.Y., Ma, F.T., Sun, P., 2019. Lactoferrin: major physiological functions and applications. Curr. Protein Pept. Sci. 20 (2), 139–144.

Jafary, F., Kasabian, S., Sharateat, S.Z.S., 2013. Purification, immobilization, and characterization of bovine lactoperoxidase. Int. J. Food Prop. 16 (4), 905–916.

Jeffrey, M.P., Jones Taggart, H., Strap, J.L., Edun, G., Green-Johnson, J.M., 2020. Milk oligosaccharides: a review. J. Dairy Sci. 103 (15), 11510–11516.

Kumar, B.S.G., Reddy, P.M., Kottekad, S., 2018. Comparative site-specific N-glycosylation analysis of chicken eggshell matrix. Food Chem. 276, 266–273.

Luo, W., Wang, J., Wang, Y., Tang, J., Ren, Y., Geng, F., 2021. Bacteriostatic effects of lactoperoxidase on Pseudomonas aeruginosa. J. Dairy Sci. 104 (10), 10528–10539.

Luo, W., Wang, J., Shi, Y., He, H., Luo, W., Huang, Q., Geng, F., 2022a. Oxalic acid is a key component involved in the early formation of egg-white gel. Food Chem. 366, 136596.

Sahoo, B.R., Maharana, J., Patra, M.C., Bhui, G.K., Lenka, S.K., Dubeys, P., Green-Johnson, J.M., 2020. Structural and dynamic investigation of bovineolkate receptor alpha (FOLR1), and role of ultra-high temperature processing on conformational and thermodynamic characteristics of FOLR1–folate complex. Colloids Surf. B Biointerfaces 121, 301–318.

Sakandar, H.A., Zhang, H.P., 2021. Trends in Probiotic(s)-Fermented milks and their influence in food safety: a review. Trends Food Sci. Technol. 110, 55–65.

Shi, P., Fan, F., Chen, H., Xu, Z., Cheng, S., Lu, W., Du, M., 2020. A bovine lactoferrin-derived peptide induced osteogenesis via regulation of osteoblast proliferation and differentiation. J. Dairy Sci. 103 (15), 11509–11516.

Singh, A.K., Singh, N., Sharma, S., Shin, K., Takase, M., Kaur, P., Srivivasan, A., Singh, S., 2019. Inhibition of lactoferrin protein by its own catalytic product: crystal structure of the hypothyroxinemia-inhibited bovine lactoferrin at 2.3-Å resolution. Biophys. J. 116 (2), 646–654.

Smoczynski, M., Staniwczek, B., Kieczewska, K., 2011. Biological activities of milk fat globule membrane proteins. Med. Wetter. 67 (5), 313–317.

Wang, J., Wen, X., Zhang, Y., Zuo, P., Cheng, L., Gan, R., Li, X., Liu, D., Geng, F., 2021a. Quantitative proteomic and metabolomic analysis of Dictyophora indusiata fruiting bodies during post-harvest morphological development. Food Chem. 339, 127848.

Wang, J., Zhao, W., Guo, S., Sun, Y., Yao, K., Liu, Z., Sun, Z., Kowk, L.Y., Peng, C., 2021b. Different growth behaviors and metabolic profiles in yoghurts induced by multistrain probiotics of Lactobacillus casei Zhang and Bifidobacterium lactis V9 under different fermentation temperatures. J. Dairy Sci. 104 (10), 10528–10539.

Wang, J., Liu, X., Li, S., He, H., Luo, W., Huang, Q., Geng, F., 2022a. Ovomucin may be the key protein involved in the early formation of egg-white gel. Food Chem. 366, 136596.

Wang, Y., Wang, J., Shi, Y., He, H., Luo, W., Geng, F., 2022b. Quantitative proteomic analysis of chicken egg yolk. Food Chem. 374, 131828.

Xiao, J., Wang, J., Cheng, L., Gao, S., Geng, F., 2020. A puzzle piece of protein N-glycosylation in chicken egg: N-glycopeptide of chicken egg vitelline membrane. Int. J. Biol. Macromol. 164 (1), 3125–3132.

Yang, R., Geng, F., Huang, X., Qiu, N., Li, S., Teng, H., Chen, L., Song, H., Huang, Q., 2020. Integrated proteomic, phosphoproteomic and N-glycoproteomic analyses of chicken eggshell matrix. Food Chem. 330, 127167.

Ye, J., 2020. Transcription factors activated through RIP (regulated intramembrane proteolysis) and RAT (regulated alternative translocation). J. Biol. Chem. 295 (30), 10271–10280.

Zhang, L., Boeren, S., van Hooijdonk, A.C.M., Vervoort, J.M., Hettinga, K.A., 2015a. A proteomic perspective on the changes in milk proteins due to high somatic cell count. J. Dairy Sci. 98 (8), 5339–5351.

Zheng, W.L., Fan, S.Q., Zou, G.Y., Shi, L., Geng, F., 2020. Underlying mechanism for the differences in heat-induced gel properties between thick egg whites and thin egg whites: gel properties, structure and quantitative proteome analysis. Food Hydrocolloids 106, 105873.

Zhong, L., Xu, L., Li, Z., Xu, S., Zhang, J., Hettinga, K., Zhou, Y., 2021a. Effects of microflotation combined with ultrasoundisation on shelf life and bioactive protein of skim milk. Ultrason. Soc. 77, 105668.

Jiao, H., Bhandari, B., Galais, C., Prakash, S., 2021. Physicochemical and microstructural properties of fermentation-induced almond emulsion-filled gels with varying concentrations of protein, fat and sugar contents. Curr. Res. Food Sci. 4, 577–584.

Zheng, J.Q., Wang, H.T., Deng, Z.F., Shan, Y.Y., Lu, X., Zhao, X., 2020. Structure and biological activities of glycoproteins and their metabolic properties in maintaining intestinal health. Crit. Rev. Food Sci. Nutr. Early Access. https://doi.org/10.1080/10408398.2021.1987857.