

Comprehensive characterization of the site-specific N-glycosylation of wild-type and recombinant human lactoferrin expressed in the milk of transgenic cloned cattle

Tian Yu^{2,*}, Chengdong Guo^{2,*}, Jianwu Wang^{2,3},
Piliang Hao⁴, Shunchao Sui², Xiuping Chen², Ran Zhang²,
Peipei Wang⁵, Guangli Yu⁵, Lei Zhang⁶, Yunping Dai^{2,6},
and Ning Li^{1,2}

²State Key Laboratory for Agrobiotechnology, China Agricultural University, Yuanmingyuan West Road 2, Beijing 100193; ³Institute of Zoology, Chinese Academy of Sciences, Beijing 100101; ⁴Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100029; ⁵School of Medicine and Pharmacy, Ocean University of China, Qingdao 266100; and ⁶Beijing GenProtein Biotech Company Ltd., Beijing 100193, People's Republic of China

Received on March 31, 2010; revised on September 17, 2010; accepted on September 18, 2010

The glycosylation profile of a recombinant protein is important because glycan moieties can play a significant role in the biological properties of the glycoprotein. Here we determined the site-specific N-glycosylation profile of human lactoferrin (hLF) and recombinant human lactoferrin (rhLF) expressed in the milk of transgenic cloned cattle. We used combined approaches of monosaccharide composition analysis, lectin blot, glycan permethylation and sequential exoglycosidase digestion and analyzed samples using high-performance ion chromatography and mass spectrometry (MS). N-glycans from hLF are comprised entirely of highly branched, highly sialylated and highly fucosylated complex-type structures, and many contain Lewis^x epitopes. Six of these structures are reported here for the first time. However, N-glycans from rhLF are of the high mannose-, hybrid- and complex-type structures, with less N-acetylneuraminic acid and fucose. Some contain a terminal N-acetylgalactosamine–N-acetylglucosamine (LacdiNAc) disaccharide sequence. Monosaccharide composition analysis of rhLF revealed small amounts of N-glycolylneuraminic acid, which were not detected by MS. hLF and rhLF appear to be glycosylated at the same two sites: Asn138 and Asn479. The third putative glycosylation site, at Asn624, is unglycosylated in both hLF and rhLF. The relative abundance of each N-glycan at each site was also determined.

The different N-glycosylation profile of rhLF when compared with that of hLF is in consistent with the widely held view that glycosylation is species- and tissue/cell-specific. These data provide an important foundation for further studies of glycan structure/function relationships for hLF and rhLF and help to better understand the glycosylation mechanism in bovine mammary epithelial cells.

Keywords: lactoferrin / N-glycan / site-specific N-glycosylation / transgenic cloned cattle

Introduction

Production of recombinant human proteins of pharmaceutical interest has been achieved using various expression systems, such as bacterial, yeast, insect, plant and mammalian cells and, more recently, transgenic animals. Use of a transgenic animal mammary bioreactor to express human proteins has become increasingly popular because of established methodologies and economically promising production processes (Niemann et al. 2005), faithful incorporation of posttranslational modifications and efficiency of purification of heterologous proteins. So far, more than 10 recombinant proteins have been produced in the milk of transgenic animals and have been shown to have similar biological properties to the native proteins. Examples include human α -glucosidase (Bijvoet et al. 1996) and human interferon- γ (James et al. 1995) in mice, α 1-antitrypsin (Massoud et al. 1991) and C1 inhibitor (Koles et al. 2004) in rabbits, human protein C (Velander et al. 1992) and human factor VIII (Paleyanda et al. 1997) in pigs, human antithrombin (Edmunds et al. 1998) and human tissue-type plasminogen activator (Denman et al. 1991; Ebert et al. 1991) in goats, and human lactoferrin (hLF; van Berkel et al. 2002) and human α -lactalbumin (Wang et al. 2008) in cows. Most of these proteins are glycoproteins, for which their glycan moieties play physiological roles. In many cases, the presence and nature of oligosaccharides may impact protein folding, protein targeting and trafficking, ligand recognition and binding, immunogenicity, pharmacological and biological activity, pharmacokinetic profile, solubility and stability (Dwek 1995; Walsh and Jefferis 2006). It is well known that protein glycosylation of recombinant proteins differs according to the specific cell lines and/or species used for expression. It is important, therefore, to determine and also take into account

¹To whom correspondence should be addressed: Tel: +86-10-62731142; Fax: +86-10-62733394; E-mail: ninglbau@public3.bta.net.cn
*These two authors contributed equally to this work.

the specific and stable glycosylation patterns of recombinant proteins in their host cells. Compared with other host cell lines, the specificity and functionality of mammary gland glycosylation has not been clearly documented.

hLF is a multifunctional, 80-kDa glycoprotein that is involved in many physiological processes, such as host defense (Nuijens et al. 1996), inflammatory regulation (Baveye et al. 1999) and growth promotion (Naot et al. 2005), as well as iron absorption in the intestinal tract (Lonnerdal and Bryant 2006). These bioactivities suggest that hLF may have important therapeutic applications in human health. Market demand for hLF is therefore primed to expand dramatically. Many attempts to produce recombinant human lactoferrin (rhLF) using prokaryotic and eukaryotic expression systems have been made, but the resulting unsuitable posttranslational modifications have limited applications (Ward et al. 1992; Legrand et al. 1995; Salmon et al. 1998; Chong and Langridge 2000). Before using rhLF as a therapeutic protein, the glycosylation profile must be elucidated to compare the different glycoforms produced in various expression systems and ultimately to ensure that this difference does not result in detrimental clinical properties.

The glycosylation patterns of hLF and rhLF produced in several expression systems have been extensively studied. The wild-type hLF sequence contains three potential N-glycosylation sites, located at Asn138, Asn479 and Asn624. The first two sites are glycosylated by complex-type N-glycans, whereas the third site (Asn624) is mostly unglycosylated (Spik et al. 1982; van Berkel et al. 1996). The glycosylation profile of rhLF expressed in transgenic maize, tobacco and rice has also been investigated (Samyn-Petit et al. 2001, 2003; Fujiyama et al. 2004). rhLF expressed in these systems contains typical plant paucimannose-type N-glycans, with β -1,2-xylose and α -1,3-linked fucose at the proximal GlcNAc. The fact that different glycosylation patterns affect protein activity is not evident, but it is potentially allergenic to humans. Only a few publications have reported on glycosylation of rhLF in transgenic animal bioreactors, but no systematic investigation has been undertaken so far (Nuijens et al. 1997; van Berkel et al. 2002; Zhang et al. 2008).

Recently, our laboratory has produced functional rhLF in the milk of transgenic cloned cattle using a novel procedure that combines gene transfer in cultured somatic cells and somatic cell nuclear transfer. A 150-kb bacterial artificial chromosome containing the entire hLF genomic sequence allowed the expression of the recombinant glycoprotein at levels >2.5 g/L in cow milk. rhLF produced using this method has very similar properties to native hLF, such as susceptibility to proteolysis, iron binding and releasing properties and antibacterial properties (Yang et al. 2008). The molecular mass of this rhLF is 2 kDa smaller than that of hLF, suggesting that the two types have different glycosylation profiles. Until now, whether the N-glycan structures affect the function of rhLF has not been clearly understood. Our aim, therefore, was to characterize the N-glycosylation of rhLF purified from the milk of transgenic cloned cattle, in order to provide information on how N-glycans are synthesized in bovine mammary epithelial cells. We also compared the site-specific N-glycosylation of hLF and rhLF to provide a foundation for further investigations into how the differential glycosylation impacts protein function.

The general strategy employed to investigate the site-specific N-glycosylation of hLF and rhLF is outlined in Figure 1. The amount of oligosaccharides and sialic acid and their distribution are determined by comparing the mass of the protein before and after deglycosylation and desialylation. The N-glycans of hLF and rhLF are identified using a combination of monosaccharide composition analysis, lectin blot, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) mapping, MALDI-TOF/TOF MS, gas chromatography/mass spectrometry (GC/MS) and sequential exoglycosidase digestions. Micro-heterogeneity at each site was determined by MS analysis of the glycopeptides and deglycosylated peptides. Using this strategy, we systematically demonstrated the site-specific N-glycosylation of rhLF purified from the milk of transgenic cloned cattle and detected eight previously unreported low-abundant glycans of hLF.

Results

MS determination of glycosylation and sialylation of hLF and rhLF

MALDI-TOF MS was used to determine the molecular mass of the proteins hLF and rhLF before and after deglycosylation. The mass of intact rhLF was 79,498 Da, slightly lower than that of hLF which was 81,074 Da (Figure 2). After deglycosylation, the observed average mass of rhLF was 76,321 Da, close to the mass of deglycosylated hLF which was 76,359 Da. This result is in good agreement with the theoretical average molecular mass of 76,320 Da calculated from the amino acid sequence of hLF and confirms that N-glycosylation is the major posttranslational modification of hLF and rhLF. The observed difference in the mass between hLF and rhLF indicates that the glycan moiety of hLF is larger (and probably more complex) than that of rhLF, comprising ~5.8% of the mass of hLF when compared with 4.0% for rhLF.

To estimate the contribution of sialylated glycans, hLF and rhLF were treated with a sialidase. After desialylation, the mass of hLF was reduced by 1.5 kDa, whereas that of rhLF was reduced by 0.5 kDa. This indicates that the sialic acid content of hLF is two times higher than that of rhLF. Thus, hLF has more sialylated glycans than rhLF, and the difference in molecular mass between intact rhLF and hLF may be due mainly to the differing sialic acid content. The lower levels of sialic acid in rhLF suggest that it may be predominantly glycosylated with neutral glycans.

Monosaccharide composition analysis of hLF and rhLF

The monosaccharide compositions of hLF and rhLF are shown in Table I. Neutral monosaccharides were analyzed using high-performance ion chromatography (HPIC), and sialic acid content was evaluated using reversed-phase high-performance liquid chromatography (RP-HPLC) after derivatization with 1,2-diamino-4,5-methylene dioxybenzene (DMB). The results are expressed relative to the proportion of Man, which was arbitrarily set at 3.0.

Both hLF and rhLF contained GlcNAc, Fuc, Man and Gal, but hLF had relatively higher amounts of GlcNAc, Fuc and Gal when compared with rhLF. The higher ratios of Fuc and

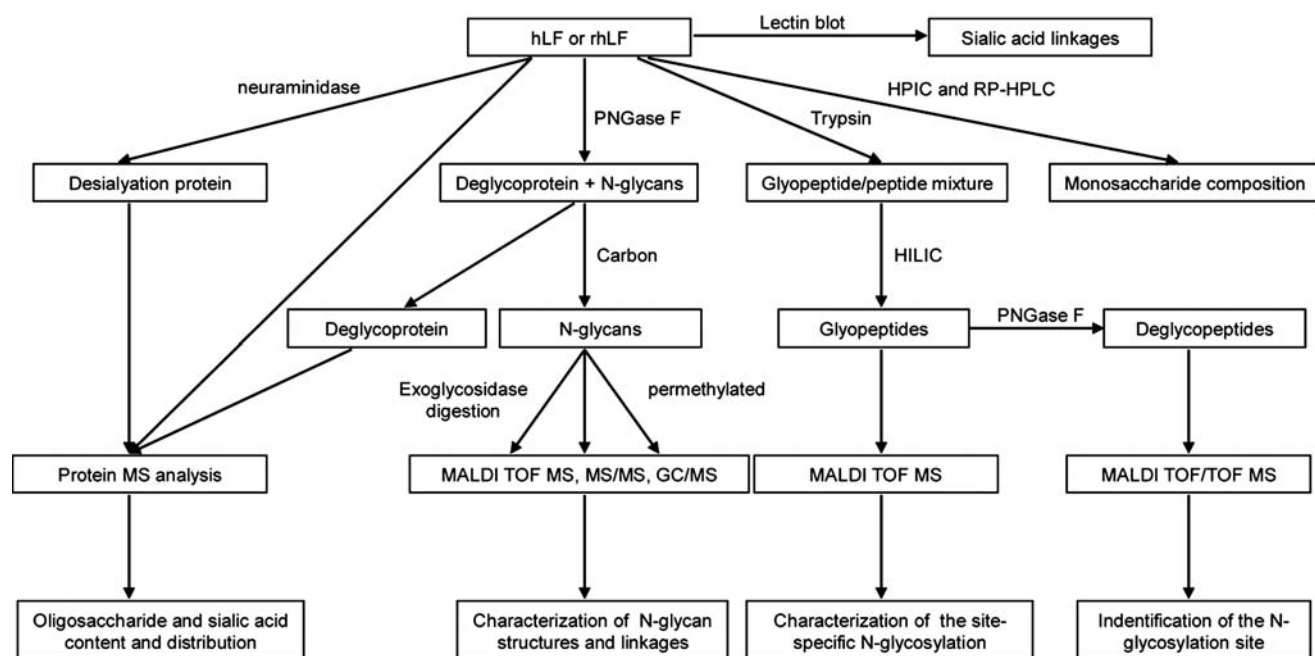


Fig. 1. Strategy for the characterization of site-specific N-glycosylation of hLF and rhLF.

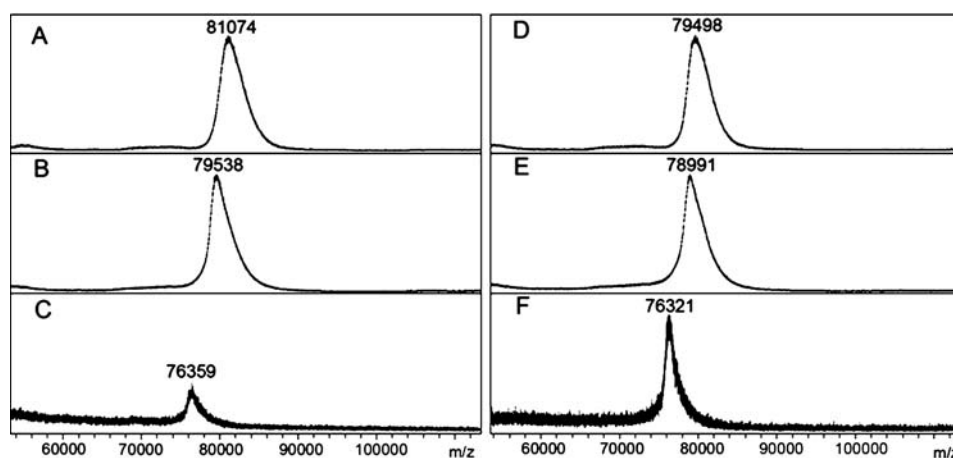


Fig. 2. Deglycosylation and desialylation of hLF and rhLF analyzed by MALDI-TOF MS. (A) Native hLF. (B) PNGase F-treated hLF. (C) Neuraminidase-treated hLF. (D) Native rhLF. (E) PNGase F-treated rhLF. (F) Neuraminidase-treated rhLF.

GlcNAc to Man suggest that the hLF glycan is highly branched and highly fucosylated. rhLF contained also GalNAc, which was not present in hLF. The ratios of GlcNAc and Gal to Man suggest that the rhLF glycan contained GalNAc next to GlcNAc and also had high mannose and/or hybrid-type structures. Additionally, the monosaccharide composition analysis shows that both hLF and rhLF contained *N*-acetylneuraminic acid (Neu5Ac) with the Neu5Ac content of hLF being 50% higher than that of rhLF. This may mean that the *N*-glycans of hLF are more highly branched and/or have more complete sialylation than those of rhLF. As anticipated, however, no *N*-glycolylneuraminic acid (Neu5Gc) was detected in the hLF glycans, although small amounts were

present in rhLF glycans. Neu5Ac levels in rhLF were six times higher than Neu5Gc levels, suggesting that rhLF may primarily be sialylated by Neu5Ac in bovine mammary epithelial cells.

Analysis of the Neu5Ac linkages of hLF and rhLF using lectin binding assays

In order to identify the Neu5Ac linkages of hLF and rhLF, we do lectin blot using digoxigenin (DIG)-labeled Maackia amurensis agglutinin (MAA) which can recognize α -2,3-linked Neu5Ac, DIG-labeled *Sambucus nigra* agglutinin (SNA) which can recognize α -2,6-linked Neu5Ac. As shown in

Table I. Monosaccharide composition of hLF and rhLF^a

Monosaccharide	hLF	rhLF
Fuc	1.8	0.3
GalNAc	—	0.9
GlcNAc	3.9	2.0
Gal	1.8	0.2
Man	3.0	3.0
Neu5Ac	0.9	0.2
Neu5Gc	—	0.03

^aThe values are means of three separate experiments.

Supplementary data, Figure S1, both hLF and rhLF have strong signals incubated with SNA, whereas very faint signals incubated with MAA. This result indicates that the sialic acid linkages of hLF and rhLF were very similar. For both proteins, Neu5Ac was predominantly α -2,6-linked to the oligosaccharide terminal antennae.

Analysis of the native N-glycan structures of hLF and rhLF using MALDI-TOF MS

To characterize the N-linked glycans of hLF and rhLF, the peptide N-glycosidase F (PNGase F)-released N-glycans were desalted and analyzed by MALDI-TOF MS. The most abundant N-glycans of hLF are complex-type glycan structures (Spik et al. 1982; Samyn-Petit et al. 2003). The terminal sialic acid residues of such structures are easily lost in the positive-ion reflectron mode, and so the negative-ion linear mode, using 2,4,6-trihydroxyacetophenone (THAP) as the matrix, was used to identify these N-glycans (Spik et al. 1982; Papac et al. 1996). The positive-ion reflectron mode was used for rhLF, which has relatively less sialic acid than does hLF and may be predominantly glycosylated by neutral glycans. Deduced N-glycan structures were obtained by comparison of observed mass to theoretical glycan mass generated using the GlycoMod tool available at ExPASy (Cooper et al. 2001). Data from the MS analysis are shown in Figure 3 and summarized in Table II.

Using the negative-ion linear mode, a series of $[M-H]^-$ adducts were observed for native hLF N-glycans (Figure 3A) at m/z values of 1931.5, 2077.6, 2223.7, 2369.0, 2442.8, 2589.5, 2735.8, 2882.1, 2954.5, 3101.0 and 3246.6.

These correspond to a number of acidic glycans comprised of biantennary and triantennary branches, with mono- or di-sialylation, and a variable extent of fucosylation. No neutral glycans were found. Most of the identified glycans contained at least one Fuc residue; several glycans contained two, three or four, indicating that in these instances, Fuc is linked not only to the core GlcNAc, but also to the antennae—a characteristic feature of Fuc within the Lewis epitope.

For rhLF, the N-glycans were comprised of a very heterogeneous mixture of oligosaccharides (Figure 3B). Using the positive-ion reflectron mode, we observed a series of $[M+Na]^+$ adducts at m/z values of 1257.37, 1419.42, 1581.47, 1743.52 and 1905.57 and these correspond to high-mannose structures ($\text{Hex}_{2-6}\text{Man}_3\text{GlcNAc}_2$). The signals at m/z values of 1501.48, 1622.50 and 1647.53 correspond to the hybrid glycan structures $\text{Hex}_1\text{HexNAc}_2\text{Man}_3\text{GlcNAc}_2$, $\text{Hex}_3\text{HexNAc}_1\text{Man}_3\text{GlcNAc}_2$ and $\text{Fuc}_1\text{Hex}_1\text{HexNAc}_2\text{Man}_3\text{GlcNAc}_2$, respectively. Signals at m/z values of 1663.52 and 1809.58 are consistent with $\text{Hex}_2\text{HexNAc}_2\text{Man}_3\text{GlcNAc}_2$ and $\text{Fuc}_1\text{Hex}_2\text{HexNAc}_2\text{Man}_3\text{GlcNAc}_2$ structures, respectively. Signals at m/z values of 1704.55, 1745.57, 1850.60 and 1891.62 are in consistent with those of complex glycan structures with neutral biantennary branches carrying one or two HexNAc residues (instead of Gal) at the nonreducing ends of the antennae. No acidic glycans were detected in the positive-ion reflectron mode, although three acidic glycans of rhLF have been observed using the negative-ion linear mode (Supplementary data, Figure S1 and Table I). All of these correspond to the monosialylated biantennary glycan structures with HexNAc-HexNAc at the nonreducing ends. No signals containing Neu5Gc were detected.

Analysis of the permethylated N-glycan structures of hLF and rhLF using MALDI-TOF MS and MALDI-TOF/TOF MS

Released glycans from hLF and rhLF were permethylated and analyzed using MS (Figure 4). The proposed identities of the permethylated glycan structures of hLF are shown in Figure 4A and Table II; this further confirms results obtained from the negative-ion linear mode MS except for the most low-abundant structure at m/z value of 3246.6 ($[M-H]^-$). A number of neutral biantennary glycans at m/z values of 2069.84, 2244.01, 2418.08, 2592.17 and 2867.28 were also detected. No theoretical masses for high-mannose or

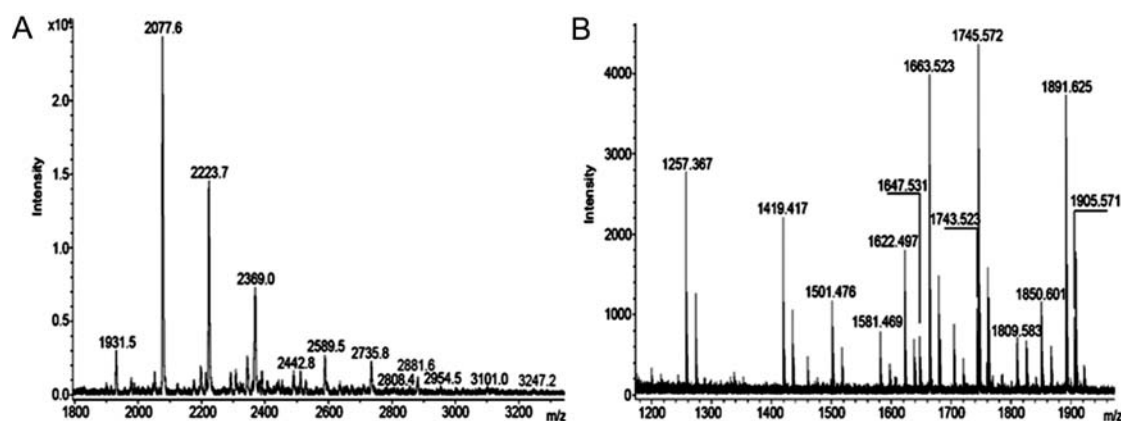


Fig. 3. MALDI-TOF MS spectra of PNGase F-released N-glycans of (A) hLF and (B) rhLF.

Table II. Proposed carbohydrate composition of N-glycans released from hLF and rhLF observed by MS in Figure 2

Protein	Native (Δm)/adduct	Permethylation (Δm)	Composition
hLF	—/[M - H] ⁻	2069.84 (-0.14)	Gal ₂ GlcNAc ₂ + Man ₃ GlcNAc ₂
	—/[M - H] ⁻	2244.01 (-0.11)	Fuc ₁ Gal ₂ GlcNAc ₂ + Man ₃ GlcNAc ₂
	—/[M - H] ⁻	2418.08 (-0.13)	Fuc ₂ Gal ₂ GlcNAc ₂ + Man ₃ GlcNAc ₂
	1931.5 (-0.26)/[M - H] ⁻	2431.08 (-0.13)	Gal ₂ GlcNAc ₂ NeuAc ₁ + Man ₃ GlcNAc ₂
	—/[M - H] ⁻	2592.17 (-0.13)	Fuc ₃ Gal ₂ GlcNAc ₂ + Man ₃ GlcNAc ₂
	2077.6 (-0.3)/[M - H] ⁻	2605.16 (-0.14)	Fuc ₁ Gal ₂ GlcNAc ₂ NeuAc ₁ + Man ₃ GlcNAc ₂
	2223.7 (-0.34)/[M - H] ⁻	2779.24 (-0.19)	Fuc ₂ Gal ₂ GlcNAc ₂ NeuAc ₁ + Man ₃ GlcNAc ₂
	—/[M - H] ⁻	2867.28 (-0.17)	Fuc ₂ Gal ₃ GlcNAc ₃ + Man ₃ GlcNAc ₂
	2369.0 (-0.16)/[M - H] ⁻	2966.31 (-0.16)	Fuc ₁ Gal ₂ GlcNAc ₂ NeuAc ₂ + Man ₃ GlcNAc ₂
	2442.8 (-0.44)/[M - H] ⁻	3054.40 (-0.13)	Fuc ₁ Gal ₃ GlcNAc ₃ NeuAc ₁ + Man ₃ GlcNAc ₂
	2589.5 (0.12)/[M - H] ⁻	3228.43 (-0.18)	Fuc ₂ Gal ₃ GlcNAc ₃ NeuAc ₁ + Man ₃ GlcNAc ₂
	2735.8 (0.27)/[M - H] ⁻	3402.50 (-0.20)	Fuc ₃ Gal ₃ GlcNAc ₃ NeuAc ₁ + Man ₃ GlcNAc ₂
	2881.6 (-0.07)/[M - H] ⁻	3576.56 (-0.17)	Fuc ₄ Gal ₃ GlcNAc ₃ NeuAc ₁ + Man ₃ GlcNAc ₂
	2808.4 (-0.17)/[M - H] ⁻	3503.49 (-0.26)	Fuc ₁ Gal ₄ GlcNAc ₄ NeuAc ₁ + Man ₃ GlcNAc ₂
	2954.7 (-0.02)/[M - H] ⁻	3677.64 (-0.20)	Fuc ₂ Gal ₄ GlcNAc ₄ NeuAc ₁ + Man ₃ GlcNAc ₂
	3101.0 (-0.14)/[M - H] ⁻	3851.68 (-0.29)	Fuc ₃ Gal ₄ GlcNAc ₄ NeuAc ₁ + Man ₃ GlcNAc ₂
	3246.6 (-0.40)/[M - H] ⁻	—	Fuc ₄ Gal ₄ GlcNAc ₄ NeuAc ₁ + Man ₃ GlcNAc ₂
rhLF	1257.37 (-0.05)/[M + Na] ⁺	1579.67 (-0.11)	Hex ₂ + Man ₃ GlcNAc ₂
	1419.42 (-0.06)/[M + Na] ⁺	1783.76 (-0.12)	Hex ₃ + Man ₃ GlcNAc ₂
	1501.48 (-0.05)/[M + Na] ⁺	1865.85 (-0.08)	Hex ₁ HexNAc ₂ + Man ₃ GlcNAc ₂
	1581.47 (-0.06)/[M + Na] ⁺	1987.85 (-0.13)	Hex ₄ + Man ₃ GlcNAc ₂
	1622.50 (-0.06)/[M + Na] ⁺	2028.88 (-0.13)	Hex ₃ HexNAc ₁ + Man ₃ GlcNAc ₂
	1647.53 (-0.05)/[M + Na] ⁺	2039.90 (-0.13)	Fuc ₁ Hex ₁ HexNAc ₂ + Man ₃ GlcNAc ₂
	1663.52 (-0.06)/[M + Na] ⁺	2069.90 (-0.13)	Hex ₂ HexNAc ₂ + Man ₃ GlcNAc ₂
	1704.55 (-0.05)/[M + Na] ⁺	2110.90 (-0.16)	Hex ₁ HexNAc ₃ + Man ₃ GlcNAc ₂
	1743.52 (-0.08)/[M + Na] ⁺	2191.95 (-0.14)	Hex ₅ + Man ₃ GlcNAc ₂
	1745.57 (-0.06)/[M + Na] ⁺	2152.00 (-0.08)	HexNAc ₄ + Man ₃ GlcNAc ₂
	1809.58 (-0.06)/[M + Na] ⁺	2244.10 (-0.02)	Fuc ₁ Hex ₂ HexNAc ₂ + Man ₃ GlcNAc ₂
	1850.60 (-0.06)/[M + Na] ⁺	2284.96 (-0.19)	Fuc ₁ Hex ₁ HexNAc ₃ + Man ₃ GlcNAc ₂
	1891.62 (-0.07)/[M + Na] ⁺	2325.98 (-0.19)	Fuc ₁ HexNAc ₄ + Man ₃ GlcNAc ₂
	1905.57 (-0.06)/[M + Na] ⁺	2396.02 (-0.16)	Hex ₆ + Man ₃ GlcNAc ₂
	2013.68 (-0.18)/[M - H] ⁻	2513.42 (0.16)	HexNAc ₄ NeuAc ₁ + Man ₃ GlcNAc ₂
	2118.70 (-0.25)/[M - H] ⁻	2646.52 (0.20)	Fuc ₁ Hex ₁ HexNAc ₃ NeuAc ₁ + Man ₃ GlcNAc ₂
	2159.79 (-0.29)/[M - H] ⁻	2687.60 (0.25)	Fuc ₁ HexNAc ₄ NeuAc ₁ + Man ₃ GlcNAc ₂

Δm is the relative mass error of the observed mass to the theoretical mass. —, not detected.

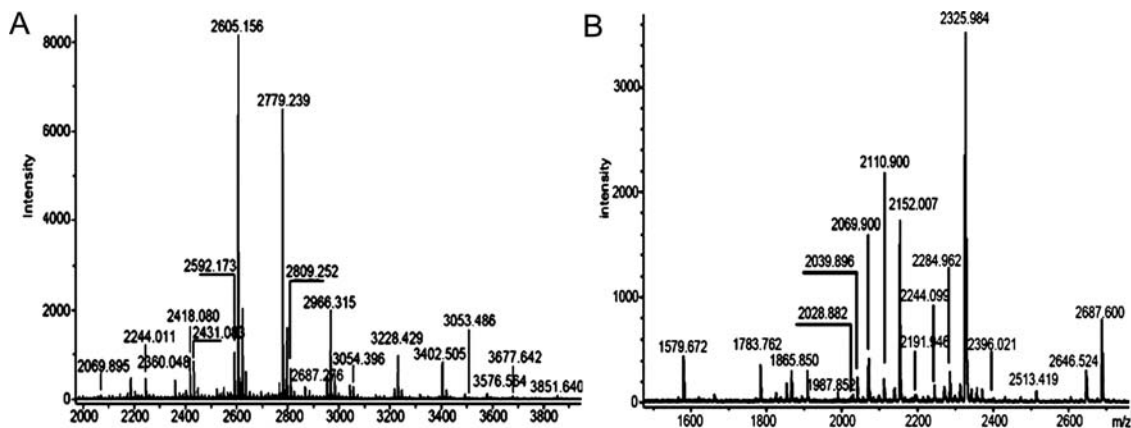


Fig. 4. MALDI-TOF MS spectra of permethylated N-glycans released from (A) hLF and (B) rhLF.

hybrid-type structures were matched to observed masses, indicating that all glycans of hLF were complex-type structures. For rhLF analysis, all neutral glycans detected in the positive-ion reflectron mode were also identified after permethylation (Figure 4B and Table II). The most abundant signals were from-neutral biantennary structures with HexNAc-HexNAc epitope. In addition, three signals at *m/z* values of 2513.6, 2646.7 and

2687.6 arising from acidic glycans were found, which corresponded to HexNAc₄NeuAc₁ + Man₃GlcNAc₂, Fuc₁Hex₁HexNAc₃NeuAc₁ + Man₃GlcNAc₂ and Fuc₁HexNAc₄NeuAc₁ + Man₃GlcNAc₂, respectively. To determine the fine structural details of the assigned molecular compositions and to facilitate the unambiguous sequence, some uncertain species of hLF and rhLF detected

were subjected to complementary modes of MS/MS analyses. For hLF, most structures could have been determined by MALDI-TOF MS except for some highly fucosylated glycans. By low-energy collision-induced dissociation (CID) on a Bruker MALDI-TOF/TOF instrument, the distribution of Fuc on terminal epitopes versus core structures could easily be defined by the abundant sodiated B and Y ions arising. Thus, structural features including Neu5Ac-sialylated Hex-HexNAc, Hex-HexNAc with one and two Fuc residues (Lewis^{x/a} and Lewis^{y/b}) and antennal LacNAc-LacNAc extensions with one, two and three Fuc residues were identified by the sodiated B ions at *m/z* values of 847, 660, 834, 1109, 1283 and 1458, respectively, and further corroborated by the Y ions arising from consecutive losses of these nonreducing terminal epitopes from the parent ions. Representative data

from analysis of *m/z* 3328, 3402 and 3576 are shown in Figure 5A–C, respectively. The MS/MS results indicated that these glycans are biantennary structures not triantennary which contain Lewis^{x/a}-LacNAc-LacNAc, Lewis^{x/a}-LacNAc-Lewis^{x/a}-LacNAc or Lewis^{y/b}-LacNAc-LacNAc at the nonreducing ends of the antennae. For rhLF analysis, MS/MS was analyzed using ABI MALDI-TOF/TOF 4800. The previous results showed that the glycans were composed of high mannose structures, hybrid structures and complex biantennary structures with HexNAc-HexNAc at the nonreducing end and core fucosylation. MS/MS results were further confirmed that and defined the structural features. Data representative of these experiments from analysis of *m/z* 1579, 1865 and 2326 are shown in Figure 5D, E and F, respectively. Signal at *m/z* value of 1579 is typical high mannose structures which

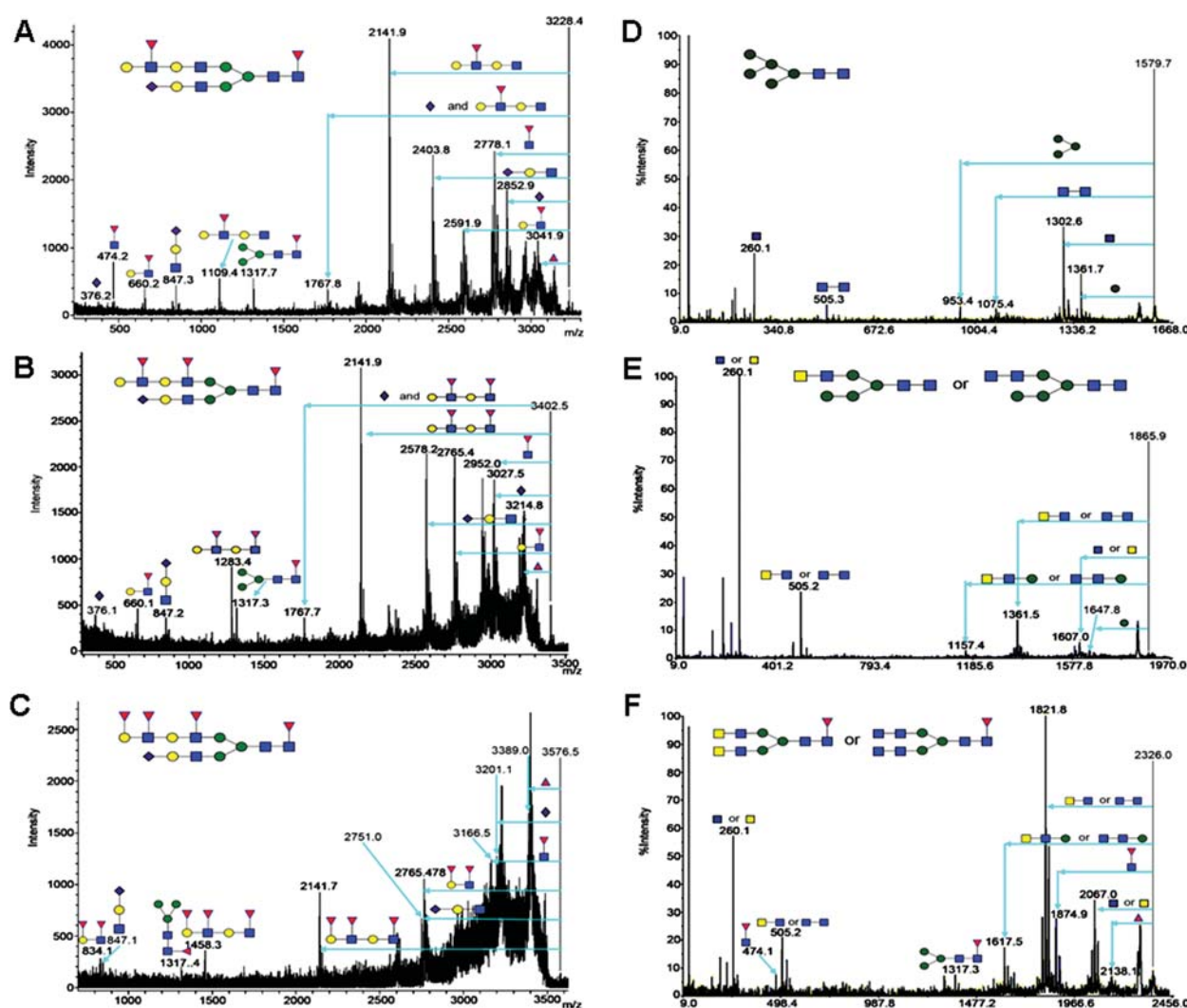


Fig. 5. MALDI-TOF/TOF mass spectrum of permethylated *N*-glycans from hLF and rhLF. (A–C) MS/MS of the molecular ion at *m/z* 3228.4 (composition Fuc₂Gal₃GlcNAc₃NeuAc₁ + Man₃GlcNAc₂), 3402.5 (composition Fuc₃Gal₃GlcNAc₃NeuAc₁ + Man₃GlcNAc₂) and 3576.5 (composition Fuc₃Gal₃GlcNAc₃NeuAc₁ + Man₃GlcNAc₂), derived from permethylated *N*-glycans from hLF (Figure 4A). (D–F) MS/MS of the molecular ion at *m/z* 1579.7 (composition Hex₂ + Man₃GlcNAc₂), 1865.9 (composition Hex₁HexNAc₂ + Man₃GlcNAc₂) and 2326.0 (composition Fuc₁HexNAc₄ + Man₃GlcNAc₂), derived from permethylated *N*-glycans from rhLF. Assignments of the fragment ions are shown. GlcNAc, blue square; GalNAc, yellow square; Man, green circle; Gal, yellow circle; Fuc, red triangle; NeuAc, purple diamond.

contains two terminal mannose at the nonreducing end. MS/MS of 1865 and 2326 show that they are hybrid structure and biantennary complex structure with core fucosylation, respectively. The present fragment ion at m/z 505 and the absent ion at m/z 486 confirm that both glycans contain HexNAc-HexNAc at antennae end instead of classic Gal-GlcNAc epitope. However, MS/MS could not distinguish LacdiNAc from GlcNAc-GlcNAc. As the data of monosaccharide compositions shown, rhLF contains more GalNAc than hLF, so it may most likely exist LacdiNAc motif.

Linkage analysis of the PNGase F-released glycans from hLF and rhLF

To identify the position of glycosidic bonds, the permethylated *N*-glycans of hLF and rhLF were analyzed by GC/MS. The data on linked glycosyl composition analysis for hLF and rhLF are presented in Figure 6A and B, respectively.

Key features of these data are as follows: (1) 2-linked mannose, 3,6-linked mannose and 4-linked GlcNAc were detected in both samples, which are typically found among *N*-linked glycans. The abundant 2-linked Man indicates that the majority of the complex glycans are biantennary but low levels of 2,6-linked Man suggest that minor triantennary structures are also present. The absence of detectable levels of 3,4,6-Man and 2,4-linked Man indicates that bisecting GlcNAc and tetraantennary structures are unlikely to be present. (2) Terminal galactose was prominent for hLF, whereas terminal mannose was pronounced for rhLF and also a strong terminal HexNAc peak was observed. This is a convincing evidence that the glycan structures of hLF are mainly of complex-type having galactose at their nonreducing end, whereas the glycans of rhLF contain more high mannose-type structures and most of the complex-type glycans were terminated on the nonreducing ends with *N*-acetylhexosamines. (3) The 6-linked galactose where sialic acid could have possibly attached was detected in both samples but the intensity was greater for hLF than rhLF. However, 6-linked GalNAc, which suggests that a sialic was attached to this residue, was detected for rhLF. This further confirmed that the sialic acid for both samples are most probably 2,6-linkage and the 6-linked

GalNAc indicated that rhLF has GlcNAc-GalNAc-NeuAc structures. (4) Terminal fucose was also present in both samples as well as in 4,6-linked GlcNAc, indicating that fucose could have been attached to the latter residue (GlcNAc) on the reducing end. Additionally, for hLF, 3,4-linked GlcNAc was detected, suggesting that fucose could have been linked to this residue elsewhere in the oligosaccharide structure but not at the reducing end. The high-level 4-linked GlcNAc indicates that terminal galactose is mainly 1,4-linked to GlcNAc. So, fucose is most likely 1,3-linked to GlcNAc. This, in addition to the MS/MS results and previous report in GlycoSuiteDB (P02788), further determined that hLF glycans contain Lewis^x and Lewis^y structures rather than the isomeric Lewis^a and Lewis^b structures.

Sequential exoglycosidase digestion of the PNGase F-released glycans

To define the anomeric configurations and to confirm the structures proposed from the initial MS profile, the PNGase F-released glycans were split into four aliquots, subjected to sequential exoglycosidase treatment and analyzed by MS (Figure 7 and Supplementary data, Tables S1 and S2). For hLF, the first aliquot of glycans was digested with neuraminidase, which cleaves α -(2-3,6,8,9)-linked neuraminic acid residues. As expected, all sialylated molecules previously described (Figure 3A and Table II) were converted into other compounds with reduced molecular weight, with the loss of one or two sialic acid residues (Figure 7A and Supplementary data, Table S1). The second aliquot of glycans was treated with neuraminidase and β -galactosidase, to confirm the complex-type oligosaccharide structures. The MS data showed that not all Gal residues were removed, because β -galactosidase can cleave β -(1-3,4,6)-linked Gal which exposed at the terminal end of the antennary provided Fuc is not bound to the subterminal GlcNAc. The nonterminal galactose that exists in LacNAc-LacNAc structures could not be removed (Figure 7B and Supplementary data, Table S1). After digestion of neuraminidase and β -galactosidase, a third aliquot of glycans was treated with β -*N*-acetylhexosaminidase (Figure 7C and Supplementary data, Table S1). This enzyme

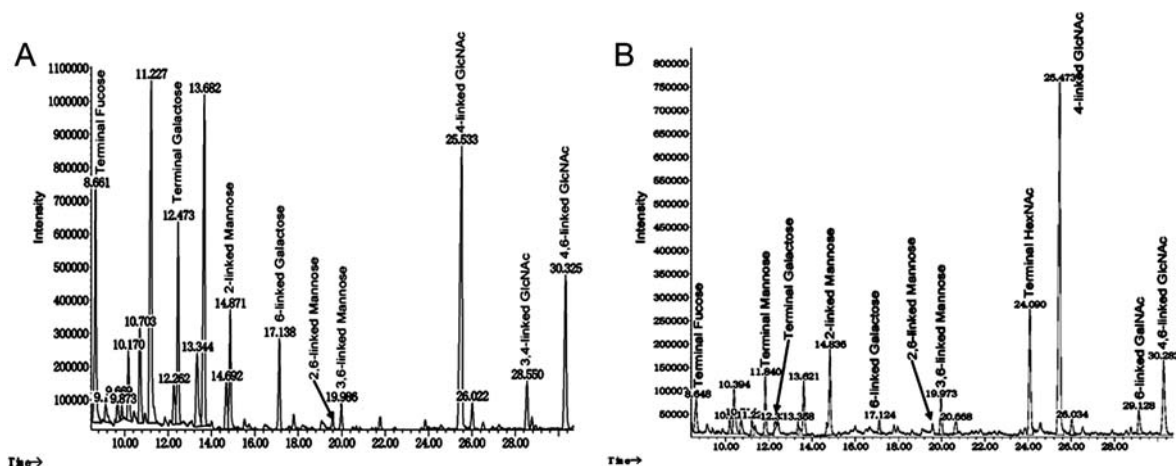


Fig. 6. GC chromatogram of PMAAs of released *N*-linked glycans from (A) hLF and (B) rhLF.

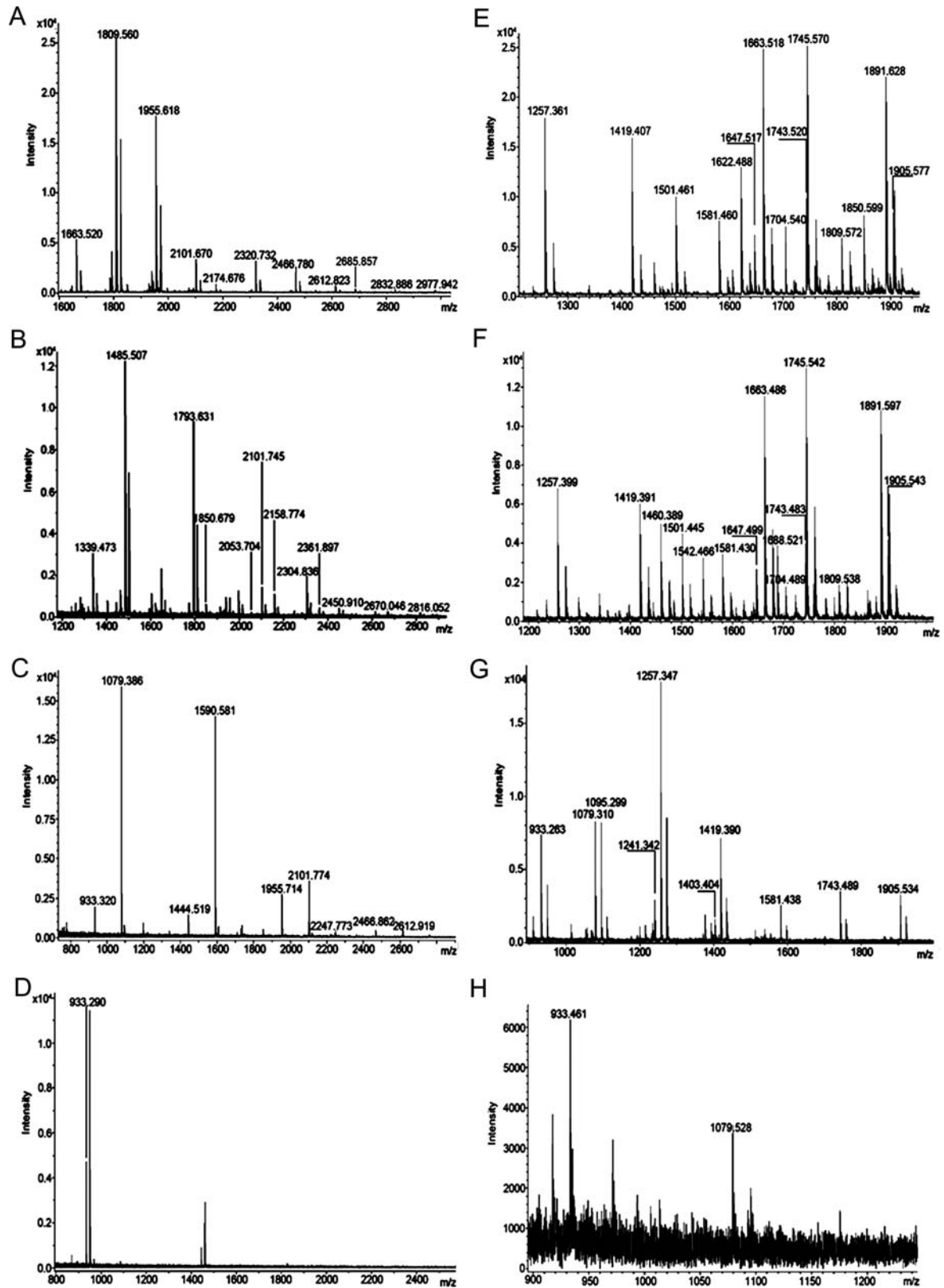


Fig. 7. MALDI-TOF MS spectra of *N*-glycans from hLF and rhLF after exoglycosidase digestions. (A) hLF and (E) rhLF after treatment with neuraminidase. (B) hLF and (F) rhLF after treatment with neuraminidase and β -galactosidase. (C) hLF and (G) rhLF after treatment with neuraminidase, β -galactosidase and β -N-acetylhexosaminidase. (D) hLF after treatment with neuraminidase, β -galactosidase, and β -N-acetylhexosaminidase and α -fucosidase. (H) rhLF after treatment with neuraminidase, β -galactosidase, β -N-acetylhexosaminidase and α -mannosidase.

removes β -(1–2,3,4,6)-bound GlcNAc residues: type I or II antennae without Fuc bound to the subterminal GlcNAc will be removed completely, whereas Lewis structures within glycan antennae structure will not be cleaved. From the MS results, some glycans were not digested completely, indicating that these glycans indeed contain Lewis structures.

The fourth aliquot of glycans was treated with neuraminidase, β -galactosidase, β -*N*-acetylhexosaminidase and α -fucosidase (Figure 7D and Supplementary data, Table S1). Only one signal at m/z 933 was observed, corresponding to the trimannosyl core (Man₃GlcNAc₂). It can be concluded, therefore, that the glycans of hLF have no high mannose- or hybrid-type structures and are composed solely of complex-type structures.

After neuraminidase treatment of the rhLF glycans, there was no obvious change in the MS spectrum (using the positive-ion reflectron mode) because the major glycans of rhLF are neutral structures. The signals from the three minor acidic glycans detected previously moved to m/z values of 1745, 1850 and 1891 which increase in the intensity, that is in consistent with the removal of neuraminic acid (Figure 7E and Supplementary data, Table S2).

After combined digestion with neuraminidase and β -galactosidase (Figure 7F and Supplementary data, Table S2), the $[M + Na]^+$ signals at m/z 1622, 1704 and 1850 were significantly reduced in the intensity, and a concomitant increase in the intensity of the molecular ions at m/z 1460 (Hex₂HexNAc₁Man₃GlcNAc₂), 1542 (HexNAc₃Man₃GlcNAc₂) and 1688 (HexNAc₃Fuc₁Man₃GlcNAc₂) was observed. This is in consistent with the removal of one or two terminal β -Gal residues.

Digestion of glycans with a combination of neuraminidase, β -galactosidase and β -*N*-acetylhexosaminidase (Figure 7G and Supplementary data, Table S2) resulted in a greatly simplified MS spectrum, characterized by major signals at m/z 933 (Man₃GlcNAc₂), 1079 (Fuc₁Man₃GlcNAc₂), 1095 (Hex₁Man₃GlcNAc₂), 1257 (Hex₂Man₃GlcNAc₂), 1419 (Hex₃Man₃GlcNAc₂), 1581 (Hex₄Man₃GlcNAc₂), 1743 (Hex₅Man₃GlcNAc₂) and 1905 (Hex₆Man₃GlcNAc₂). These signals were not affected by the foregoing exoglycosidase digestions, which indicate that these glycans are high-mannose structures. Two minor signals at m/z 1241 (Fuc₁Hex₁Man₃GlcNAc₂) and 1403 (Fuc₁Hex₂Man₃GlcNAc₂) were also observed. Molecular ions at m/z values of 1542 (HexNAc₃Man₃GlcNAc₂), 1688 (Fuc₁HexNAc₃Man₃GlcNAc₂), 1745 (HexNAc₄Man₃GlcNAc₂) and 1891 (Fuc₁HexNAc₄Man₃GlcNAc₂) were efficiently digested by β -*N*-acetylhexosaminidase to products at m/z 1079 and 933, which is in consistent with a trimannosyl core with and without core fucosylation, respectively. This result indicates that these glycans are complex-type oligosaccharide structures. The disappearance of signals at m/z 1501 (Hex₁HexNAc₂Man₃GlcNAc₂), 1647 (Hex₁HexNAc₂Fuc₁Man₃GlcNAc₂), 1663 (Hex₂HexNAc₂Man₃GlcNAc₂) and 1809 (Fuc₁Hex₂HexNAc₂Man₃GlcNAc₂) is accompanied by a concomitant increase in the intensity of molecular ions at m/z 1095 (Hex₁Man₃GlcNAc₂), 1257 (Hex₂Man₃GlcNAc₂) and 1403 (Fuc₁Hex₂Man₃GlcNAc₂), which is in consistent with the composition of these being hybrid structures. After

α -mannosidase treatment (Figure 7H and Supplementary data, Table S2), the molecular ions at m/z 1095 (Hex₁Man₃GlcNAc₂), 1403 (Fuc₁Hex₂Man₃GlcNAc₂), 1257 (Hex₂Man₃GlcNAc₂), 1419 (Hex₃Man₃GlcNAc₂), 1581 (Hex₄Man₃GlcNAc₂), 1743 (Hex₅Man₃GlcNAc₂) and 1905 (Hex₆Man₃GlcNAc₂) were all “trimmed” to m/z 933 (Man₃GlcNAc₂) and 1079 (Fuc₁Man₃GlcNAc₂). This result further confirms that these structures were of the high-mannose type.

According to the monosaccharide composition analysis, linkage analysis and MALDI-TOF MS results, rhLF may have LacdiNAc epitopes at the nonreducing end which is very hard to distinguish from GlcNAc-GlcNAc by MS/MS. In order to determine the nonreducing structures of rhLF, we analyzed the permethylated glycans on the MALDI after *N*-acetyl- β -D-glucosaminidase (GKX-80050) and β -*N*-acetyl-hexosaminidase (GKX-5003), successively. *N*-acetyl- β -D-glucosaminidase can specifically cleave all non-reducing β -linked *N*-acetylglucosamines, and after that N-linked glycans are the same as those in the intact (untreated Figure 4B) sample (Figure 8A). However, after additional treatment of β -*N*-acetyl-hexosaminidase (jack bean), all terminal *N*-acetylhexosamines on the nonreducing end were removed effectively (Figure 8B). This result indicates that GalNAc was the outmost *N*-acetylhexosamine and after its release, it was followed by the sequential removal of GlcNAc. So, it further confirmed that the *N*-glycans of rhLF have terminal LacdiNAc disaccharide sequence.

Through monosaccharide composition analysis, MALDI-TOF MS, MS/MS and GC/MS as well as sequential exoglycosidase digestion, the compositions and the linkages of the *N*-glycans from hLF and rhLF are summarized in Figure 9. The *N*-glycans of hLF and rhLF identified are significantly different. hLF glycans are comprised entirely of complex-type oligosaccharides that are more complex, more sialylated and more fucosylated than the glycans of rhLF. Most of these contain Lewis^x structures, whereas H13 also contains a Lewis^y structure. Six glycans (H12–H17) composed of biantennary and triantennary structures containing the LacNAc-LacNAc epitope, with monosialylation and varying degrees of fucosylation were first detected on hLF. The glycans of rhLF, on the other hand, contain a mix of high mannose, hybrid and complex-type oligosaccharide structures. Most of these are neutral glycans, and some contain the LacdiNAc disaccharide epitope. Only three sialylated glycans were found however, and we were unable to find evidence for *N*-glycan structures terminated with Neu5Gc.

Identification of *N*-glycosylation sites

hLF contains three putative *N*-glycosylation sites at Asn138, Asn479 and Asn624 (Rey et al. 1990; van Berkel et al. 1996). Theoretical masses of the predicted tryptic peptides containing the putative glycosylation sites are summarized in Table III. To identify the *N*-glycan site occupancy on hLF and rhLF, the glycopeptides were selectively enriched using hydrophilic interaction chromatography (HILIC) microcolumns and analyzed in the positive linear mode by MALDI-TOF MS (Figure 10A and B). Two well separated peak clusters were observed and were assigned to the glycopeptides. Heterogeneity of the carbohydrate moiety is seen by mass

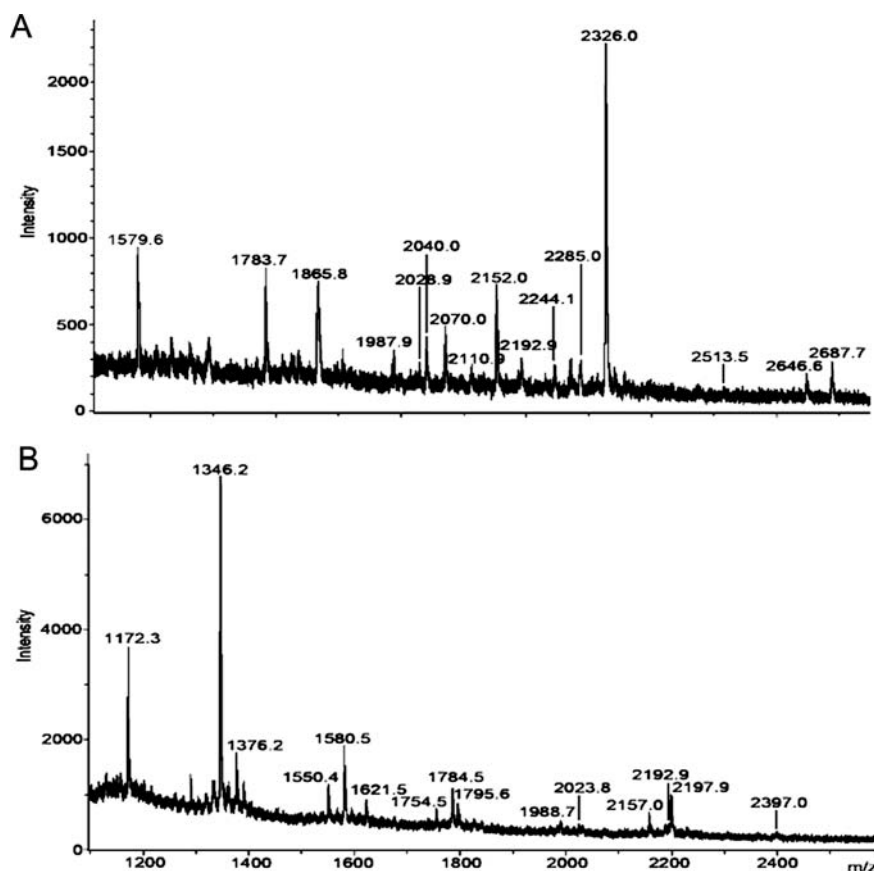


Fig. 8. MALDI-TOF MS of PNGase F-released permethylated N-linked glycans from rhLF after digestion with special exoglycosidases. (A) After treatment with β -D-glucosaminidase (GKX-80050). (B) After treatment with β -D-glucosaminidase (GKX-80050) and β -N-acetylhexosaminidase (jack bean GKX-5003).

differences of 146, 162 and 291 Da, corresponding to the loss of fucose, hexose and neuraminic acid residues, respectively. The major mass differences resulted from the loss of fucose and neuraminic acid for hLF and from the loss of mannose for rhLF. Following *N*-glycosidase F treatment of the glycopeptides, two MS peaks at *m/z* 2096 and 3231 were observed for both hLF and rhLF (Figure 10C and D), corresponding to peptides containing the N-glycosylation sites Asn138 and Asn479, respectively. The peptide sequences of these were further confirmed by CID MS/MS and matched by searching the Mascot database (Figure 11). No glycopeptide containing site Asn624 was detected in either hLF or rhLF, which suggests that this site is not glycosylated.

Heterogeneity of the N-glycans at each glycosylation site

After elucidating the structures of the released *N*-glycans and mapping the glycopeptides, the microheterogeneity of the *N*-glycans at each glycosylation site was investigated. We wanted to determine whether specific glycan structures dominated at each glycosylation site. By subtracting the theoretical mass of the tryptic peptides from the experimentally observed mass of the glycopeptides, information on the oligosaccharide residues can be obtained. The previously identified glycans were matched, one to one, to the glycopeptides. The relative abundance of the glycoforms at each glycosylation site was

estimated from the MS signal strength, giving a site-specific glycoprofile (Table IV).

Glycopeptides from hLF generally had a higher mass than those from rhLF, and the glycans were all complex-type structures. At site Asn479, hLF glycans were all biantennary structures with or without Fuc and sialic acid, whereas at Asn138 the glycans were more complex, more branched and more fucosylated. The two predominant glycopeptides that contain Asn479 and Asn138 have *m/z* values of 4156.9 and 5293.0, respectively, and the corresponding glycan moiety could be a mix of biantennary monosialylated species (H5) and neutral biantennary trifucosylated species (H6). As the molecular masses of these two glycans differ only by 1 Da, based on the MS data of glycopeptides, they could not be distinguished. Eight minor signals are observed at *m/z* values of 5658.1, 5803.9, 5950.2, 6024.8, 6096.3, 6170.3, 6316.3 and 6462.0. This is in consistent with the biantennary and triantennary monosialylated species (H10–H17) that are attached to Asn138.

For rhLF, the glycans are all neutral structures at site Asn479; the two predominant signals at *m/z* values of 3313.6 and 3475.5 correspond to glycopeptides containing the high-mannose species R1 and R2. At site Asn138, however, the glycoforms are mostly complex neutral biantennary structures; the major signal at *m/z* 5082.8 corresponds to a glycopeptide containing the glycan species R13, which has two GalNAc

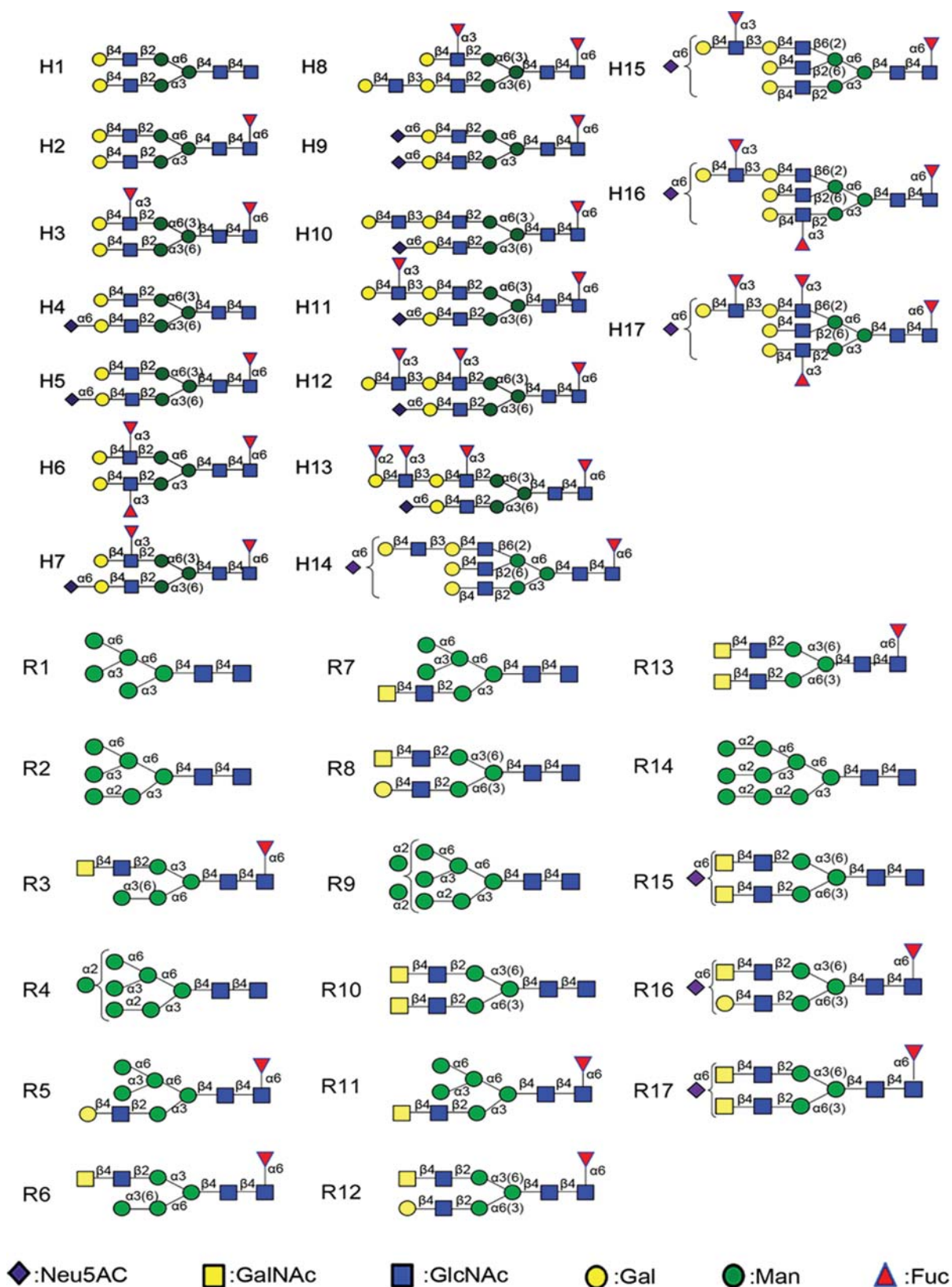


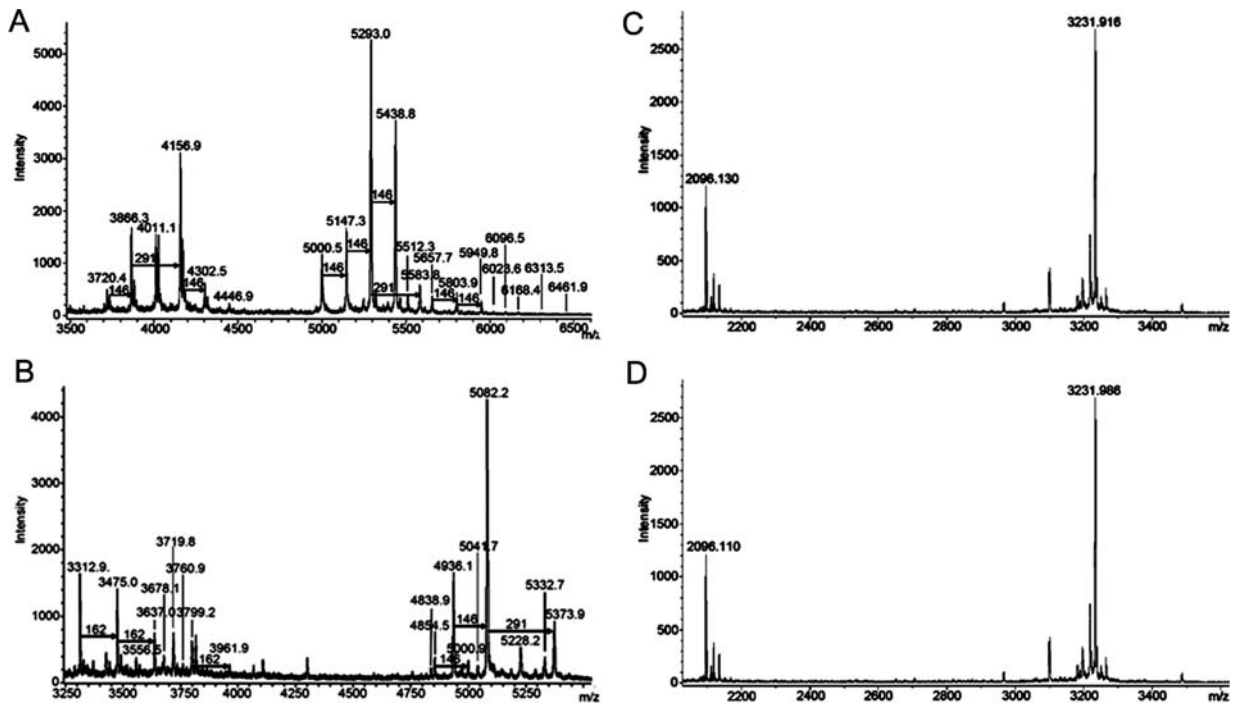
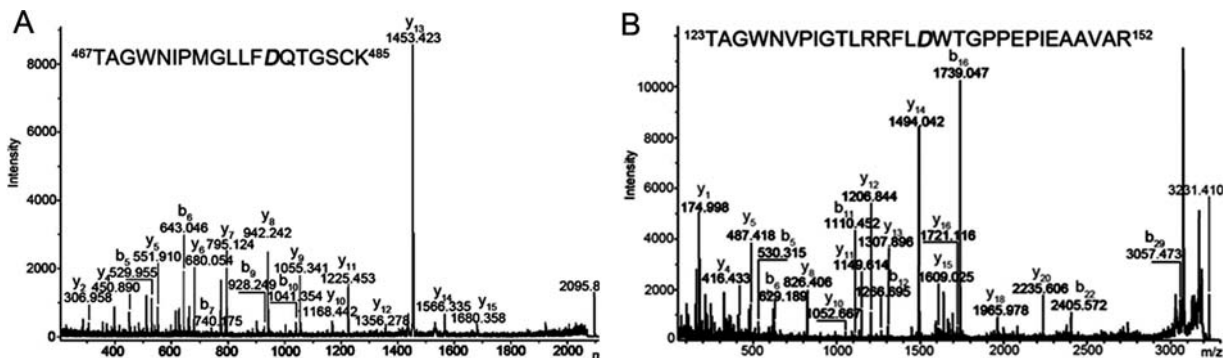
Fig. 9. The *N*-glycan structures identified for hLF (H1–H17) and rhLF (R1–R17) are cartooned according to MS-Tools from EUROcarbDB. GlcNAc, blue square; GalNAc, yellow square; Man, green circle; Gal, yellow circle; Fuc, red triangle; NeuAc, purple diamond.

Table III. Tryptic peptides containing putative N-glycosylation sites for hLF

Theoretical mass	Amino acid	Peptide sequence ^a	Mass after deglycosylation ^b	Mass after alkylation ^c
3230.72	123–152	TAGWNVPIGTLRPFL N WTGPPEPIEA ^a AVAR	3231.71	
2037.99	467–485	TAGWNIPMGLLF N QTG SCK	2038.99	2096.01
835.32	624–631	NGSDCPDK	836.32	892.35

^aPutative N-glycosylation sites are shown in boldface.

^bAfter deglycosylation, Asn (N) is converted to Asp (D), increasing peptide mass by 1 Da.

^cAfter alkylation, peptides containing Cys (C) are carboxamidomethylated, increasing peptide mass by 57 Da.

Fig. 10. MALDI-TOF MS spectra of the glycopeptides and deglycopeptides from hLF and rhLF. hLF and rhLF before treatment (A, B) and after treatment (C, D) with PNGase F. The two glycopeptides are corresponded to contain the sites Asn479 and Asn138, respectively. Mass differences corresponding to the monosaccharides are as follows: 146, fucose; 162, hexose; 291, Neu5Ac.

Fig. 11. CID-MS/MS spectra of the parent ions at (A) m/z 2095.80 and (B) m/z 3231.39, corresponding to the peptides containing the sites Asn479 and Asn138, respectively. The identities of the b and y fragment ions are indicated above the fragment ion masses. After PNGase F treatment, N479 and N138 are converted to D479 and D138, respectively.

residues at the nonreducing ends and a core Fuc residue. Three acidic glycopeptides at m/z values of 5228.11, 5333.01 and 5374.45 were also detected. These correspond to

glycopeptides glycosylated with the glycan moieties R15, R16 and R17, respectively. All of these are terminated with Neu5Ac. No glycopeptides containing Neu5Gc were

Table IV. The N-glycoforms present on each site of hLF and rhLF and their relative abundance

N-glycosylation site	Adduct	m/z		Glycoforms	Relative abundance (%) ^a
		Experimental	Theoretical		
hLF Asn479	[M + H] ⁺	3720.4	3719.9	H1	6.28 ± 0.25
	[M + H] ⁺	3866.4	3866.1	H2	18.78 ± 0.51
	[M + H] ⁺	4011.1	4011.2	H4	20.34 ± 0.71
	[M + H] ⁺	4156.9	4157.3	H5	38.77 ± 0.28 ^b
	[M + H] ⁺	4156.9	4158.4	H6	
	[M + H] ⁺	4302.5	4303.5	H7	10.41 ± 0.93
	[M + H] ⁺	4446.9	4448.6	H9	4.79 ± 0.19
hLF Asn138	[M + H] ⁺	5000.5	5002.3	H2	4.51 ± 0.25
	[M + H] ⁺	5147.3	5148.5	H3	7.50 ± 1.00
	[M + H] ⁺	5293.0	5293.6	H5	32.37 ± 0.80 ^b
	[M + H] ⁺	5293.0	5294.6	H6	
	[M + H] ⁺	5438.8	5439.7	H7	27.46 ± 1.30
	[M + H] ⁺	5512.3	5513.8	H8	2.44 ± 0.11
	[M + H] ⁺	5583.8	5584.9	H9	7.73 ± 0.30
	[M + H] ⁺	5657.7	5658.9	H10	3.52 ± 0.29
	[M + H] ⁺	5803.9	5805.1	H11	5.28 ± 0.07
	[M + H] ⁺	5949.8	5951.2	H12	4.19 ± 0.02
	[M + H] ⁺	6023.6	6024.3	H14	1.07 ± 0.12
	[M + H] ⁺	6096.5	6097.4	H13	1.45 ± 0.07
	[M + H] ⁺	6168.4	6170.4	H15	1.09 ± 0.11
	[M + H] ⁺	6313.5	6316.6	H16	0.80 ± 0.01
	[M + H] ⁺	6461.9	6462.7	H17	0.60 ± 0.02
rhLF Asn479	[M + H] ⁺	3312.9	3313.5	R1	29.79 ± 0.40
	[M + H] ⁺	3475.0	3475.7	R2	21.94 ± 0.72
	[M + H] ⁺	3556.5	3557.8	R3	6.77 ± 0.42
	[M + H] ⁺	3637.0	3637.8	R4	11.39 ± 0.19
	[M + H] ⁺	3678.1	3678.9	R5	4.81 ± 0.28
	[M + H] ⁺	3719.8	3719.9	R7	8.49 ± 0.70
	[M + H] ⁺	3760.9	3761.0	R8	2.16 ± 0.21
	[M + H] ⁺	3799.2	3799.9	R9	10.80 ± 0.38
	[M + H] ⁺	3961.9	3962.1	R14	3.84 ± 0.24
	[M + H] ⁺	4838.9	4840.2	R6	4.08 ± 0.11
rhLF Asn138	[M + H] ⁺	4854.5	4856.2	R7	4.73 ± 0.12
	[M + H] ⁺	4936.1	4938.3	R10	15.94 ± 0.10
	[M + H] ⁺	5000.9	5002.3	R11	4.55 ± 0.16
	[M + H] ⁺	5041.7	5043.4	R12	3.10 ± 0.19
	[M + H] ⁺	5082.2	5084.4	R13	47.45 ± 0.77
	[M + H] ⁺	5228.2	5229.6	R15	4.27 ± 0.32
	[M + H] ⁺	5332.7	5334.7	R16	4.89 ± 0.37
	[M + H] ⁺	5373.9	5375.7	R17	10.98 ± 0.26

^aThe relative abundance (%) at each site was evaluated by the means of three times using the mass signals.
^bThe most abundant glycopeptide at each site of hLF, which probably linked to two glycoforms H5 and H6 that could not be distinguished on the mass.

observed. In summary, the glycoforms at each glycosylation site for hLF are very different from those for rhLF, but for both hLF and rhLF the glycan profile is more complex at Asn138 when compared with that at Asn479.

Discussion

The objective of this work was to analyze the site-specific N-glycosylation of rhLF expressed in the milk of transgenic cloned cattle and to compare it with that of native hLF. Although information on the glycosylation of rhLF from the milk of a transgenic cow has been previously reported, only its monosaccharide composition, rather than site-specific characterization of N-glycosylation, was analyzed (van Berkel et al. 2002). More significantly, we used a somatic cell clone to produce transgenic animals, which is different from the traditional pronuclear microinjection approach because somatic

cell reprogramming occurs and protein glycosylation may be changed during this process. The glycosylation of rhLF expressed in transgenic cloned cattle therefore warrants investigation, to ensure the safety of the process. To date, however, there are only a few reports available on the glycosylation of recombinant proteins produced in the mammary glands of transgenic cloned animals.

We have shown that N-glycosylation is the major posttranslational modification of hLF and rhLF and that the difference in mass between hLF and rhLF is due to the different N-linked glycans attached to each protein. We analyzed the N-glycan structures of hLF and rhLF using MS, MS/MS, monosaccharide composition analysis, lectin blot, GC/MS and sequential exoglycosidase digestions. In total, 17 different glycoforms were found in hLF, including the first reported six glycoforms that were found in relatively low abundance. We attribute this to the use of a THAP matrix for the MALDI-TOF MS, which enables a lower limit of detection to

be achieved (Papac et al. 1996). The observed glycans are common biantennary and triantennary complex-type structures, with differing degrees of fucosylation and some of which contain Lewis^x-LacNAc, Lewis^x-LacNAc-LacNAc and Lewis^x-LacNAc-Lewis^x-LacNAc structures. Compared with hLF, rhLF has relatively less Neu5Ac, Gal, GlcNAc and Fuc content and more Man, indicating that the *N*-glycans of rhLF are less branched, less sialylated and less fucosylated. The major *N*-glycans of rhLF are of three types: high mannose, hybrid and complex-type structures, with or without core fucosylation. Also, the charged *N*-glycans were mainly α -2,6-monosialylated with Neu5Ac, which was also the same with hLF. In combined with MS/MS, GC/MS and specific exoglycosidase digestions, a major nonreducing epitope, LacdiNAc, was found in the complex- or hybrid-type glycans, instead of the classical Gal-GlcNAc residues. This structure is also found in bovine lactoferrin (bLF), bovine CD36 and in the milk of other species (Coddeville et al. 1992; Nakata et al. 1993; Montesino et al. 2008).

The identified *N*-glycan structures of rhLF are very similar to the earlier reported ones for bovine milk fat globule membrane glycoprotein especially as bovine CD36, but not very same with bLF. The glycans of bovine milk fat globule membrane glycoprotein also contain high mannose-, hybrid- and complex-type structures. Most of them have the same compositions with rhLF (Nakata et al. 1993; Sato et al. 1993). The carbohydrate moieties of bLF consist only of complex- and high-mannose-type glycans, and no hybrid-type glycan is found. The complex glycans are composed of monosialylated or disialylated biantennary structures with a terminal *N*-acetylglucosamine and core fucosylation, monosialylated biantennary structures with one terminal GalNAc residue at the nonreducing end and also a biantennary structure with Neu5Ac-LacdiNAc and α -1,3-Gal-Gal residue in the terminal nonreducing position of each antennary structure (Spik et al. 1988; Coddeville et al. 1992). rhLF glycans also have Neu5Ac-LacdiNAc but no α -1,3-Gal-Gal residue detected. These similarities and differences indicate that rhLF takes bovine glycosylation style in the bovine mammary gland, but protein glycosylation also has protein sequence specificity that the glycans decorated may be related to the rhLF sequence, rhLF fold style and the speed of rhLF through the ER.

Location of the *N*-glycosylation sites was determined, and site-specific glycan microheterogeneity was characterized. rhLF has the same *N*-glycosylation sites as hLF, located at Asn138 and Asn479. The third putative glycosylation site at Asn624 was not glycosylated in either hLF or rhLF. This result is consistent with previous reports that glycosylation at site Asn624 might be limited by proximity to glycosylation at site Asn479 (van Berkel et al. 1996; Samyn-Petit et al. 2003).

Glycopeptides were selectively concentrated using HILIC microcolumns, which in turn significantly intensify MS signals by reducing interference from nonglycopeptide species. Site-specific glycosylation profiles were obtained by matching the identified glycans to the corresponding glycopeptide. We also used MS signal intensity to quantify the oligosaccharides at each glycosylation site. This correlates well with the relative abundances of the various glycoforms (Thaysen-Andersen et al. 2009). The most abundant

glycoforms at each site of hLF are likely composed of a biantennary monosialylated structure with one Fuc (H5) and a neutral biantennary tri-fucosylated structure (H6). In rhLF, the predominant glycoforms at Asn479 are high mannose-type structures, and at Asn138, neutral biantennary structures with core fucosylation and structures with an LacdiNAc disaccharide epitope. These results indicate that although the rhLF is correctly synthesized in bovine mammary gland and the same occupancy of glycosylation sites is retained, its specific glycosylation pattern distinguishes it from the wild-type hLF. For both hLF and rhLF, glycosylation at Asn138 resulted in more complex glycan structures than at Asn479. This difference in the relative distribution of glycoforms between the two *N*-glycosylation sites may be related to their relative positions in the polypeptide chain, and it has been suggested that glycosylation sites closer to the N terminus have greater complexity in their glycan moiety (Rudd and Dwek 1997; Fenaille et al. 2007). Perhaps, Asn138 is the first glycosylation site encountered in the Golgi apparatus and is subjected to more complete glycosylation by glycosyltransferases.

On the basis of these observations, we speculate that the differences in glycosylation patterns between hLF and rhLF may be caused mainly by different glycosidase/glycosyltransferase systems or enzymatic activities in human and bovine mammary epithelial cells. hLF contains more complex, more sialylated and more fucosylated glycans with Lewis^x and Lewis^x-LacNAc-LacNAc structures. In bovine milk, the high mannose- and hybrid-type oligosaccharides are most likely generated by the limiting action of Golgi α -mannosidase II, and the Gal residue of LacNAc unit on many glycoproteins is replaced by a GalNAc resulting in an LacdiNAc structure which cannot be formed to Lewis structures or extended to diLacNAc structures. This is also emerged on rhLF. The relatively high abundance of glycans containing the LacdiNAc epitope may be associated with *N*-acetylgalactosaminyltransferases acting in competition with a β -1,4-galactosyltransferase on the terminal GlcNAc-containing oligosaccharides in milk or may be attributable to α -lactalbumin, which is responsible for regulating β -1,4-galactosyltransferase in lactose synthesis (Montesino et al. 2008). α -Lactalbumin can interact with β -1,4-galactosyltransferases and change their conformation so that uridine diphosphate-Gal is added to glucose instead of the terminal GlcNAc (Takase and Hagiwara 1998) or to exploit uridine diphosphate-GalNAc in transferring GalNAc to biantennary acceptor glycopeptides (Do et al. 1995). The synthesis of oligosaccharides in human mammary gland is catalysed by a variety of specific glycosyltransferases, which is not dependent on α -lactalbumin (Urashima et al. 2001).

The effect of the rhLF *N*-glycans on protein function is yet to be determined. There are relatively few publications on the effect of *N*-glycans on hLF function and these remain controversial. Davidson and Lonnerdal (1988) showed that the removal of Fuc from hLF glycans significantly decreases its ability to bind the small-intestine receptor of rhesus monkeys, but it is not known how fucosylation affects *in vivo* properties. Legrand et al. (1990) showed that deglycosylation of hLF destroys protein structure and affects iron binding activity. van Berkel et al. (1995, 2002), on the other hand, showed that both glycosylated and unglycosylated hLF can bind iron and have

identical affinities for human lysozyme and bacterial lipopolysaccharide but differ in their susceptibilities to tryptic proteolysis. This suggested that glycosylation protects against proteolysis, however subsequently, that conclusion was refuted (van Veen et al. 2004). rhLF expressed in rice or in transgenic cow has also been reported as having similar bioactivities to hLF (van Berkel et al. 2002; Suzuki et al. 2003), and different glycosylation apparently did not affect protein function. But most of the experiments have been performed *in vitro*, which does not rule out other important *in vivo* roles of glycans, such as mediation of receptor function, cell adhesion and signal transduction, as well as the significant role that degree of sialylation has on pharmacokinetic properties (Bhatia and Mukhopadhyay 1999; Sinclair and Elliott 2005). For example, after desialylation, desialylated glycoproteins are recognized by the hepatic asialoglycoprotein receptors and are cleared from blood circulation quickly (Hoermann et al. 1993; Joiasse et al. 2000). As rhLF expressed in the milk of transgenic cloned cattle has relatively low levels of sialic acid and high levels of mannose as well as the LacdiNAc glycans, it may have faster clearance and therefore reduced efficacy *in vivo*.

The glycan moiety of a glycoprotein can also affect its immunogenicity by being part of the antigenic determinant. Recombinant proteins, expressed in different expression systems, have different glycosylation profiles. Potential antigenic carbohydrates in recombinant glycoproteins are Neu5Gc, α -Gal, polylactosaminic chains and blood group antigens (Jenkins and Curling 1994). The various glycan structures must be carefully analyzed for safety considerations of recombinant glycoproteins. Our results show that only a small amount of Neu5Gc was detected by HPLC, although no Neu5Gc was detected at any glycosylation site by MS. The reason for this could be that glycans sialylated with Neu5Gc were present in amounts below the detection limits of the instrument or that the small amount of Neu5Gc detected by HPLC did not arise from rhLF but was a contaminant from another glycoprotein. No other antigenic carbohydrates were found. Whether immunogenicity is affected by the different glycosylation patterns of rhLF when compared with hLF should be investigated further.

In summary, our current work provides information on the site-specific N-glycosylation of hLF and rhLF (expressed in transgenic cloned cattle) and compares their differences. The identified N-glycans are in accordance with structures found in GlycoSuiteDB and some of them are first detected. We have also characterized the monosaccharide linkages and anomery. These data are really very important for further studying the relationship between the rhLF N-glycan structure and *in vivo* function and are also useful for optimizing the glycosylation of rhLF through glycoengineering, so it will be closer to the human profile. The glycans of rhLF reflect the specific mechanism of protein glycosylation in bovine mammary epithelial cells, which is regulated by specific bovine glycosyltransferase systems. For large-scale production of more humanized bovine milk to improve infant formulas and more properly glycosylated therapeutic glycoproteins using the bovine mammary bioreactor, we may be able to change the glycosylation pathway or transform the glycosyltransferase systems by gene transfer or gene knock-out in

bovine mammary epithelial cells. Recently, our group has produced transgenic cloned cattle expressing human α -1,3/4-fucosyltransferases in mammary epithelial cells to synthesize the human Lewis a or Lewis b structure in milk, thereby inhibiting infection with *Helicobacter pylori*. Protein glycosylation is a so complex process, however, that we are far from understanding and perfecting glycosylation systems in, for example, animal mammary cells. With increased understanding of the glycobiology involved, and with further development of the emerging area of glycoengineering, modification of animal mammary bioreactors for the large-scale production of humanized milk and therapeutic proteins with human-like glycosylation may be possible.

Materials and methods

Materials

rhLF was expressed and purified as described (Yang et al. 2008). hLF obtained from human milk (in powder form, cell-culture tested), Proteomics Grade PNGase F (EC 3.5.1.52) from *Elizabethkingia meningoseptica*, Proteomics Grade α -(2-3, 6,8,9) neuraminidase, jack bean α -mannosidase (EC 3.2.1.24), bovine testes β -galactosidase (EC 3.2.1.23), jack bean β -N-acetylhexosaminidase (EC 3.2.1.30) and bovine kidney α -fucosidase (EC 3.2.1.51) were purchased from Sigma Chemicals (St. Louis, MO). 15 U/mL of jack bean β -N-acetylhexosaminidase (GKX-5003) and 4 U/mL of N-acetyl- β -D-glucosaminidase, recombinant from *Streptococcus pneumoniae*, expressed in *Escherichia coli* (GKX-80050), were purchased from (Prozyme, Hayward, CA). Sequencing-grade modified trypsin (EC 3.4.21.4) was purchased from Promega (Madison, WI). POROS R2 material (20- μ m bead size) was purchased from Applied Biosystems (Framingham, MA), zwitterionic HILIC material (ZICTM-HILIC, 200 Å, 10 μ m) was from SeQuant AB (Umea, Sweden) and Carbograph graphitized carbon columns (150 mg) were purchased from Alltech (Deerfield, IL). The MALDI matrices α -cyano-4-hydroxycinnamic acid (CHCA), 2,5-dihydroxybenzoic acid (DHB) and THAP were purchased from Fluka (Steinheim, Germany). Standard sugars, dithiothreitol, iodoacetamide and all other reagents were purchased from Sigma unless otherwise described.

Deglycosylation and desialylation of hLF and rhLF

Denatured glycoprotein solution (1 mg/mL) was prepared by adding 50 μ g of hLF or rhLF to 45 μ L of 20 mM ammonium bicarbonate, pH 8.3, and adding 5 μ L of denaturation solution (0.2% sodium dodecyl sulfate (SDS)/100 mM 2-mercaptoethanol). Each solution was heated to 100°C for 5 min to denature the glycoproteins. After cooling, 5 μ L of PNGase F (500 U/mL) was added, and the mixture was left overnight at 37°C. Samples were then desalted and concentrated to 10–20 μ L using 0.5-mL, 10-kDa Microcon YM-10 Centrifugal Filters (Millipore, Billerica, MA). The deglycosylated protein fractions were collected and analyzed by MS. The flow-through containing the released N-glycans was desalted with nonporous graphitized microcolumns and collected for further analysis using MS. To estimate sialic acid

content, 50 µg each of hLF and rhLF were desialylated in 50 mM sodium phosphate buffer (pH 6.0) containing neuraminidase overnight at 37°C. Desialylated hLF and rhLF were analyzed by MS and compared with native hLF and rhLF to estimate the extent of sialylation.

Monosaccharide composition analysis using HPIC

hLF or rhLF glycoprotein (1 mg) was hydrolyzed in 400 µL of 2 M trifluoroacetic acid (TFA) at 110°C for 4 h to release the neutral sugars. After the reaction, samples were washed with methanol three times to remove TFA and dried in a microcentrifuge concentrator (SpeedVac, Thermo Scientific, Milford, MA). The dried samples were re-dissolved in 200 µL of ultrapure water and analyzed using an ICS-3000 HPIC instrument (Dionex, Bannockburn, IL) equipped with a 3 × 150-mm CarboPac PA-20 column fitted with a 3 × 50-mm CarboPac PA guard column (Dionex) and isocratic elution with 10 mM NaOH at a flow rate of 0.5 mL/min. The eluants were detected using a pulsed amperometric detector equipped with an electrochemical cell containing a gold working electrode and an Ag/AgCl reference electrode. For calibration purposes, a standard solution containing equimolar amounts of the monosaccharides *N*-acetylgalactosamine (GalNAc), *N*-acetylglucosamine (GlcNAc), mannose (Man), fucose (Fuc) and galactose (Gal) was prepared. The solution was serially diluted in water, and the samples were run on the HPIC instrument to produce calibration curves for calibrating the monosaccharide molar ratio. Linearity of the calibration curves was validated with a correlation coefficient greater than 0.99.

Analysis of sialic acid content of hLF and rhLF

Sialic acids were released from the glycoproteins hLF and rhLF and derivatized with DMB. Briefly, 1 mg of glycoprotein was hydrolyzed with 0.1 M TFA for 1 h at 50°C, followed by the addition of 7 mM DMB solution in 1.4 M aqueous acetic acid, containing 18 mM sodium hydrosulfite and 1 M β-mercaptoethanol. The mixture was incubated for 2 h at 60°C in the dark. After cooling on ice, 20 µL of the sample was analyzed by RP-HPLC using a 4.6 × 150-mm analytical ZORBAX C-18 column (Agilent, Santa Clara, CA) and excitation and emission wavelengths of 373 and 448 nm, respectively, for fluorescence detection. The sample was eluted using a methanol/acetonitrile (ACN)/0.05% TFA buffer (6:4:90, v/v/v) at a flow rate of 1 mL/min.

Identification of the oligosaccharide using lectin binding assays

The “DIG Glycans Differentiation Kit” (Roche, Mannheim, Germany) was used to identify specific carbohydrate motives. After separation on 10% SDS–polyacrylamide gel electrophoresis (PAGE), the glycoproteins (2 µg each) were bound to Immobilon membranes (Millipore). The membranes were incubated at 25°C for 1 h with specific lectins that were digoxigenin-labeled. Afterward, the filters were immersed in a solution containing polyclonal antidigoxigenin antibody conjugated with alkaline phosphatase for another 1 h. The staining reaction was carried out in nitro blue tetrazolium/X-phosphate solution without shaking.

Permethylation of PNGase F-released N-glycans

Glycans released from hLF with PNGase F were lyophilized in a glass vial, dissolved in dimethyl sulfoxide and permethylated using the sodium hydroxide procedure by adding powdered sodium hydride and methyl iodide into the carbohydrate solution (Ciucanu and Kerek 1984). After derivatization, the reaction products were extracted in chloroform and repeatedly washed with water. The permethylated glycans were dried and redissolved in 10 µL of methanol prior to MS analysis.

Linkage analysis of the PNGase F-released glycans from hLF and rhLF

For the determination of linked glycosyls, partially methylated alditol acetates (PMAAs) were prepared from the permethylated N-linked glycans released from hLF and rhLF. Briefly, partially permethylated glycans were hydrolyzed with HCl: water:acetic acid (0.5:1.5:8, v/v/v) at 80°C overnight, followed by reduction with NaBD₄. After hydrolysis and reduction steps, the free hydroxyls of the partially methylated alditols were acetylated with acetic anhydride:pyridine (1:1, v/v) at 100°C for 1 h to produce PMAAs.

The PMAAs were analyzed on a Hewlett Packard 5890 GC interfaced to a 5970 MSD (mass selective detector, electron impact ionization mode). The separation was performed on a 30 m EC 1 bonded, phase-fused silica capillary column (Alltech, Deerfield, IL). Electron impact mass spectra were obtained under the following conditions: oven temperature, 140°C (2.0°C/min) → 220°C (20°C/min) → 300°C (7.5 min); detector temperature, 280°C; inlet temperature, 250°C.

Sequential exoglycosidase digestions

Lyophilized hLF and rhLF *N*-glycans released by PNGase F were dissolved in 40 µL of 50 mM sodium phosphate, pH 5.0, and split into four aliquots. Each aliquot was incubated overnight at 37°C with one of four exoglycosidases or exoglycosidase mixtures. hLF aliquots were incubated with (1) neuraminidase (0.5 U); (2) neuraminidase (0.5 U) and β-galactosidase (10 mU); (3) neuraminidase (0.5 U), β-galactosidase (10 mU), β-*N*-acetylhexosaminidase (0.2 U); (4) neuraminidase (0.5 U), β-galactosidase (10 mU), β-*N*-acetylhexosaminidase (0.2 U) and α-*L*-fucosidase (0.1 U). rhLF aliquots were incubated with (1) neuraminidase (0.5 U); (2) neuraminidase and β-galactosidase (10 mU); (3) neuraminidase (0.5 U), β-galactosidase (10 mU), β-*N*-acetylhexosaminidase (0.2 U); (4) neuraminidase (0.5 U), β-galactosidase (10 mU), β-*N*-acetylhexosaminidase (0.2 U), α-mannosidase (0.5 U). After each digestion, samples were purified on the graphitized carbon microcolumns and divided into aliquots for MS analysis.

For LacdiNAc analysis, another rhLF aliquot was dissolved with 50 mM sodium citrate buffer, treated with *N*-acetyl-β-D-glucosaminidase at a concentration of 4 U/mL (GKX-80050) and incubated at 37°C. After overnight incubation, a small volume of the enzyme digest was pipetted into a screw-cap tube, dried, permethylated and analyzed by MALDI-MS to evaluate the effect of the enzyme. The remaining enzyme digest was subsequently treated with 15 U/mL of β-*N*-acetylhexosaminidase (GKX-5003). The combined

enzyme treatment was carried out also at 37°C overnight. After the second enzyme incubation, the entire remaining digest was dried, permethylated and analyzed by MALDI-MS.

In-gel tryptic digestion and extraction of tryptic peptides

Glycoprotein (20 µg) was denatured in loading buffer (25 mM Tris, pH 6.8/0.5 M DTT/0.1% SDS/0.5% bromophenol blue/50% glycerol) at 100°C for 5 min and analyzed using 10% SDS-PAGE. The gel was fixed and stained with Coomassie Brilliant Blue G-250 dye for 2 h and then was partially destained with a methanol/acetic acid/water solution (25:7:68, v/v/v) overnight.

Coomassie-stained hLF and rhLF bands were excised from the gel and transferred to 0.5-mL microcentrifuge tubes. The bands were rinsed three times with ddH₂O and cut into small pieces, which were further incubated with destaining solution (50 mM ammonium bicarbonate/50% ACN) for 20 min at room temperature several times until the color was removed. The gel pieces were then shrunk in ACN for 20 min and dried in a SpeedVac centrifuge. The separated proteins contained in the gel fragments were then reduced and alkylated by adding 50 mM ammonium bicarbonate containing 20 mM dithiothreitol and incubating for 1 h at 56°C. The buffer was replaced with 110 mM iodoacetamide. Gel fragments were allowed to react for 45 min at room temperature in the dark and then washed several times with 25 mM ammonium bicarbonate (for 20 min), dehydrated with ACN and then dried in a SpeedVac centrifuge. Finally, the dried gel pieces were incubated at 37°C overnight with 25 mM ammonium bicarbonate solution containing 1 mM CaCl₂ and trypsin in an enzyme-to-substrate ratio of 1:50 (w/w). Enzymatic digestion was terminated by the addition of 1% TFA. The supernatant was removed to a clean tube, and the peptides were further extracted from the gel pieces by rinsing once with 0.1% TFA and once with 50% ACN/0.1% TFA for 20 min. Combined extractions and supernatant were dried in a SpeedVac centrifuge.

Separation of the glycopeptides using HILIC

For selective enrichment of the glycopeptides, HILIC microcolumns were prepared by packing ZICTM-HILIC material into GELoader tips (Eppendorf, Hamburg, Germany). The dried peptide pools, derived from digested glycoproteins, were dissolved in 20 µL of washing buffer (80% ACN/2% formic acid). The peptides were applied to the HILIC microcolumn, which had been equilibrated with 50 µL of washing buffer. The column was washed with 30 µL of washing buffer, and then peptides were eluted with 10 µL of 2% formic acid. Eluant was collected, and a 0.5-µL aliquot was used for MS analysis.

PNGase F digestion of the glycopeptides and purification of the deglycosylated peptides

A 5-µL aliquot of the glycopeptide sample was transferred to a new tube and dried in a SpeedVac centrifuge, followed by the addition of 20 µL of 20 mM ammonium bicarbonate. The mixture was incubated with 1 µL PNGase F (50 U/mL) at 37°C

overnight. After digestion, the products were purified on a POROS R2 microcolumn to separate the deglycosylated peptides from the *N*-glycans. Deglycosylated peptides were eluted in 10 µL of 80% ACN/0.1% TFA. To identify the glycosylation sites, the deglycosylated peptides were further analyzed by MS.

MS analysis of glycosylated peptides, deglycosylated peptides and carbohydrate moieties of hLF and rhLF

MS spectra of the glycoproteins, glycopeptides, deglycosylated peptides, intact glycans and permethylated and exoglycosidase-digested glycans were obtained using an Ultraflex MALDI-TOF/TOF MS instrument (Bruker Daltoniks GmbH, Bremen, Germany) equipped with a nitrogen laser and ABI MALDI-TOF/TOF 4800. Samples were mixed with the corresponding matrix solution (<1 µL), applied to the MALDI target and allowed to dry at room temperature. The matrix solutions were SA (10 mg/mL) in ACN/water/TFA (50:50:0.1, v/v/v) for protein analysis; CHCA (20 mg/mL) in ACN/water/TFA (70:30:0.1, v/v/v) for peptide analysis; DHB (10 mg/mL) in ACN/water (50:50, v/v) for glycopeptide analysis; and THAP (20 mg/mL)/ammonium citrate (20 mM) in ACN/water (50:50, v/v) for native glycan and permethylated glycan analyses. Protein and glycopeptide fractions were analyzed in the linear positive-ion mode; sialylated *N*-glycans in the linear negative-ion mode and peptides; and neutral *N*-glycans and permethylated glycans in the reflector positive-ion mode. MS/MS analysis was performed using the LIFT-MS/MS facility.

Supplementary data

Supplementary data for this article is available online at <http://glycob.oxfordjournals.org/>.

Funding

This work was supported by grants from the State High-Tech Research Development Program of China: 2007AA100501, 2008ZX08007001 and 2008ZX08014002.

Acknowledgements

We thank Dr. Ji Dong Feng and the Hua Da Proteomics Laboratory in the Beijing Institute of Genomics, Chinese Academy of Sciences, for use of the MALDI-TOF MS instrumentation. We thank Dr. Parastoo in Complex Carbohydrate Research Center for helping with the linkage analysis.

Conflict of interest statement

None declared.

Abbreviations

ACN, acetonitrile; bLF, bovine lactoferrin; CHCA, α -cyano-4-hydroxycinnamic acid; CID, collision-induced dissociation; DHB, dihydroxybenzoic acid; DIG, digoxigenin; DMB, 1,2-diamino-4,5-methylene dioxybenzene; Fuc, fucose; Gal, galactose; GalNAc, *N*-acetylgalactosamine; GC/MS, gas

chromatography/mass spectrometry; GlcNAc, *N*-acetylglucosamine; HILIC, hydrophilic interaction chromatography; hLF, human lactoferrin; HPIC, high-performance ion chromatography; MAA, Maackia amurensis agglutinin; MALDI-TOF MS, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry; Man, mannose; MS, mass spectrometry; MS/MS, tandem mass spectrometry; Neu5Ac, *N*-acetylneuraminic acid; Neu5Gc, *N*-glycolylneuraminic acid; PAGE, polyacrylamide gel electrophoresis; PMAA, partially methylated alditol acetate; PNGase F, peptide *N*-glycosidase F; rhLF, recombinant human lactoferrin; RP-HPLC, reversed-phase high-performance liquid chromatography; SDS, sodium dodecyl sulfate; SNA, *Sambucus nigra* agglutinin; TFA, trifluoroacetic acid; THAP, 2,4,6-trihydroxyacetophenone; TOF, time-of-flight.

References

- Baveye S, Ellass E, Mazurier J, Spik G, Legrand D. 1999. Lactoferrin: a multifunctional glycoprotein involved in the modulation of the inflammatory process. *Clin Chem Lab Med*. 37(3):281–286.
- Bhatia PK, Mukhopadhyay A. 1999. Protein glycosylation: implications for in vivo functions and therapeutic applications. *Adv Biochem Eng Biotechnol*. 64:155–201.
- Bijvoet AG, Kroos MA, Pieper FR, de Boer HA, Reuser AJ, van der Ploeg AT, Verbeet MP. 1996. Expression of cDNA-encoded human acid alpha-glucosidase in milk of transgenic mice. *Biochim Biophys Acta*. 1308(2):93–96.
- Chong DK, Langridge WH. 2000. Expression of full-length bioactive antimicrobial human lactoferrin in potato plants. *Transgenic Res*. 9(1):71–78.
- Ciucanu I, Kerek F. 1984. A simple and rapid method for the permethylation of carbohydrates. *Carbohydr Res*. 131(2):209–217.
- Codeville B, Strecker G, Wieruszkeski JM, Vliegthart JF, van Halbeek H, Peter-Katalinic J, Egge H, Spik G. 1992. Heterogeneity of bovine lactoferrin glycans: characterization of α -D-Galp-(1 $\rightarrow$ 3)- β -D-Gal- and α -NeuAc-(2 $\rightarrow$ 6)- β -D-GalpNAc-(1 $\rightarrow$ 4)- β -D-GlcNAc-substituted N-linked glycans. *Carbohydr Res*. 236:145–164.
- Cooper CA, Gasteiger E, Packer NH. 2001. GlycoMod—a software tool for determining glycosylation compositions from mass spectrometric data. *Proteomics*. 1(2):340–349.
- Davidson LA, Lonnerdal B. 1988. Specific binding of lactoferrin to brush-border membrane: ontogeny and effect of glycan chain. *Am J Physiol*. 254(4 Pt 1):G580–G585.
- Denman J, Hayes M, O'Day C, Edmunds T, Bartlett C, Hirani S, Ebert KM, Gordon K, McPherson JM. 1991. Transgenic expression of a variant of human tissue-type plasminogen activator in goat milk: purification and characterization of the recombinant enzyme. *Nat Biotechnol*. 9(9):839–843.
- Do KY, Do SI, Cummings RD. 1995. α -Lactalbumin induces bovine milk β -1,4-galactosyltransferase to utilize UDP-GalNAc. *J Biol Chem*. 270(31):18447–18451.
- Dwek RA. 1995. Glycobiology: “towards understanding the function of sugars”. *Biochem Soc Trans*. 23(1):1–25.
- Ebert KM, Selgrath JP, DiTullio P, Denman J, Smith TE, Memon MA, Schindler JE, Monastersky GM, Vitale JA, Gordon K. 1991. Transgenic production of a variant of human tissue-type plasminogen activator in goat milk: generation of transgenic goats and analysis of expression. *Nat Biotechnol*. 9(9):835–838.
- Edmunds T, Van Patten SM, Pollock J, et al. 1998. Transgenically produced human antithrombin: structural and functional comparison to human plasma-derived antithrombin. *Blood*. 91(12):4561–4571.
- Fenaille F, Le Mignon M, Groseil C, Ramon C, Riande S, Siret L, Bihoreau N. 2007. Site-specific N-glycan characterization of human complement factor H. *Glycobiology*. 17(9):932–944.
- Fujiyama K, Sakai Y, Misaki R, Yanagihara I, Honda T, Anzai H, Seki T. 2004. N-linked glycan structures of human lactoferrin produced by transgenic rice. *Biosci Biotechnol Biochem*. 68(12):2565–2570.
- Hoermann R, Kubota K, Amir SM. 1993. Role of subunit sialic acid in hepatic binding, plasma survival rate, and in vivo thyrotropic activity of human chorionic gonadotropin. *Thyroid*. 3(1):41–47.
- James DC, Freedman RB, Hoare M, Ogonah OW, Rooney BC, Larionov OA, Dobrovolsky VN, Lagutin OV, Jenkins N. 1995. N-glycosylation of recombinant human interferon- γ produced in different animal expression systems. *Nat Biotechnol*. 13(6):592–596.
- Jenkins N, Curling EM. 1994. Glycosylation of recombinant proteins: problems and prospects. *Enzyme Microb Technol*. 16(5):354–364.
- Joziasse DH, Lee RT, Lee YC, Biessen EA, Schiphorst WE, Koeleman CA, van den Eijnden DH. 2000. α 3-Galactosylated glycoproteins can bind to the hepatic asialoglycoprotein receptor. *Eur J Biochem*. 267(21):6501–6508.
- Koles K, van Berkel PH, Pieper FR, Nuijens JH, Mannesse ML, Vliegthart JF, Kamerling JP. 2004. N- and O-glycans of recombinant human C₁ inhibitor expressed in the milk of transgenic rabbits. *Glycobiology*. 14(1):51–64.
- Legrand D, Mazurier J, Colavizza D, Montreuil J, Spik G. 1990. Properties of the iron-binding site of the N-terminal lobe of human and bovine lactoferrins: importance of the glycan moiety and of the non-covalent interactions between the N- and C-terminal lobes in the stability of the iron-binding site. *Biochem J*. 266(2):575–581.
- Legrand D, Salmon V, Codeville B, Benaissa M, Plancke Y, Spik G. 1995. Structural determination of two N-linked glycans isolated from recombinant human lactoferrin expressed in BHK cells. *FEBS Lett*. 365(1):57–60.
- Lonnerdal B, Bryant A. 2006. Absorption of iron from recombinant human lactoferrin in young US women. *Am J Clin Nutr*. 83(2):305–309.
- Massoud M, Bischoff R, Dalemans W, Pointu H, Attal J, Schultz H, Clesse D, Stinnakre MG, Pavirani A, Houdebine LM. 1991. Expression of active recombinant human α 1-antitrypsin in transgenic rabbits. *J Biotechnol*. 18(3):193–203.
- Montesino R, Toledo JR, Sanchez O, et al. 2008. Monosialylated biantennary N-glycoforms containing GalNAc–GlcNAc antennae predominate when human EPO is expressed in goat milk. *Arch Biochem Biophys*. 470(2):163–175.
- Nakata N, Furukawa K, Greenwalt DE, Sato T, Kobata A. 1993. Structural study of the sugar chains of CD36 purified from bovine mammary epithelial cells: occurrence of novel hybrid-type sugar chains containing the Neu5Ac α 2 $\rightarrow$ 6GalNAc β 1 $\rightarrow$ 4GlcNAc and the Man α 1 $\rightarrow$ 2Man α 1 $\rightarrow$ 3Man α 1 $\rightarrow$ 6Man groups. *Biochemistry*. 32(16):4369–4383.
- Naot D, Grey A, Reid IR, Cornish J. 2005. Lactoferrin—a novel bone growth factor. *Clin Med Res*. 3(2):93–101.
- Niemann H, Kues W, Carnwath JW. 2005. Transgenic farm animals: present and future. *Rev Sci Tech*. 24(1):285–298.
- Nuijens JH, van Berkel PH, Geerts ME, Hartevelt PP, de Boer HA, van Veen HA, Pieper FR. 1997. Characterization of recombinant human lactoferrin secreted in milk of transgenic mice. *J Biol Chem*. 272(13):8802–8807.
- Nuijens JH, van Berkel PH, Schanbacher FL. 1996. Structure and biological actions of lactoferrin. *J Mammary Gland Biol Neoplasia*. 1(3):285–295.
- Paleyanda RK, Velandar WH, Lee TK, Scandella DH, Gwazdauskas FC, Knight JW, Hoyer LW, Drohan WN, Lubon H. 1997. Transgenic pigs produce functional human factor VIII in milk. *Nat Biotechnol*. 15(10):971–975.
- Papac DI, Wong A, Jones AJ. 1996. Analysis of acidic oligosaccharides and glycopeptides by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Anal Chem*. 68(18):3215–3223.
- Rey MW, Woloshuk SL, deBoer HA, Pieper FR. 1990. Complete nucleotide sequence of human mammary gland lactoferrin. *Nucleic Acids Res*. 18(17):5288.
- Rudd PM, Dwek RA. 1997. Glycosylation: heterogeneity and the 3D structure of proteins. *Crit Rev Biochem Mol Biol*. 32(1):1–100.
- Salmon V, Legrand D, Slomianny MC, et al. 1998. Production of human lactoferrin in transgenic tobacco plants. *Protein Expr Purif*. 13(1):127–35.
- Samyn-Petit B, Gruber V, Flahaut C, Wajda-Dubos JP, Farrer S, Pons A, Desmaizeres G, Slomianny MC, Theisen M, Delannoy P. 2001. N-glycosylation potential of maize: the human lactoferrin used as a model. *Glycoconj J*. 18(7):519–527.
- Samyn-Petit B, Wajda Dubos JP, Chirat F, Codeville B, Demaizieres G, Farrer S, Slomianny MC, Theisen M, Delannoy P. 2003. Comparative analysis of the site-specific N-glycosylation of human lactoferrin produced in maize and tobacco plants. *Eur J Biochem*. 270(15):3235–3242.
- Sato T, Furukawa K, Greenwalt DE, Kobata A. 1993. Most bovine milk fat globule membrane glycoproteins contain asparagine-linked sugar chains with GalNAc β 1 $\rightarrow$ 4GlcNAc groups. *J Biochem*. 114(6):890–900.
- Sinclair AM, Elliott S. 2005. Glycoengineering: the effect of glycosylation on the properties of therapeutic proteins. *J Pharm Sci*. 94(8):1626–1635.

- Spik G, Coddeville B, Montreuil J. 1988. Comparative study of the primary structures of sero-, lacto- and ovotransferrin glycans from different species. *Biochimie*. 70(11):1459–1469.
- Spik G, Strecker G, Fournet B, Bouquelet S, Montreuil J, Dorland L, van Halbeek H, Vliegthart JF. 1982. Primary structure of the glycans from human lactotransferrin. *Eur J Biochem*. 121(2):413–419.
- Suzuki YA, Kelleher SL, Yalda D, Wu L, Huang J, Huang N, Lonnerdal B. 2003. Expression, characterization, and biologic activity of recombinant human lactoferrin in rice. *J Pediatr Gastroenterol Nutr*. 36(2):190–199.
- Takase K, Hagiwara K. 1998. Expression of human α -lactalbumin in transgenic tobacco. *J Biochem*. 123(3):440–444.
- Thaysen-Andersen M, Mysling S, Hojrup P. 2009. Site-specific glycoprofiling of N-linked glycopeptides using MALDI-TOF MS: strong correlation between signal strength and glycoform quantities. *Anal Chem*. 81(10):3933–3943.
- Urashima T, Saito T, Nakamura T, Messer M. 2001. Oligosaccharides of milk and colostrum in non-human mammals. *Glycoconj J*. 18(5):357–371.
- van Berkel PH, Geerts ME, van Veen HA, Kooiman PM, Pieper FR, de Boer HA, Nuijens JH. 1995. Glycosylated and unglycosylated human lactoferrins both bind iron and show identical affinities towards human lysozyme and bacterial lipopolysaccharide, but differ in their susceptibilities towards tryptic proteolysis. *Biochem J*. 312(Pt 1):107–114.
- van Berkel PH, van Veen HA, Geerts ME, de Boer HA, Nuijens JH. 1996. Heterogeneity in utilization of N-glycosylation sites Asn624 and Asn138 in human lactoferrin: a study with glycosylation-site mutants. *Biochem J*. 319(Pt 1):117–122.
- van Berkel PH, Welling MM, Geerts M, van Veen HA, Ravensbergen B, Salaheddine M, Pauwels EK, Pieper F, Nuijens JH, Nibbering PH. 2002. Large scale production of recombinant human lactoferrin in the milk of transgenic cows. *Nat Biotechnol*. 20(5):484–487.
- van Veen HA, Geerts ME, van Berkel PH, Nuijens JH. 2004. The role of N-linked glycosylation in the protection of human and bovine lactoferrin against tryptic proteolysis. *Eur J Biochem*. 271(4):678–684.
- Velander WH, Johnson JL, Page RL, Russell CG, Subramanian A, Wilkins TD, Gwazdauskas FC, Pittius C, Drohan WN. 1992. High-level expression of a heterologous protein in the milk of transgenic swine using the cDNA encoding human protein C. *Proc Natl Acad Sci USA*. 89(24):12003–12007.
- Walsh G, Jefferis R. 2006. Post-translational modifications in the context of therapeutic proteins. *Nat Biotechnol*. 24(10):1241–1252.
- Wang J, Yang P, Tang B, et al. 2008. Expression and characterization of bioactive recombinant human α -lactalbumin in the milk of transgenic cloned cows. *J Dairy Sci*. 91(12):4466–4476.
- Ward PP, May GS, Headon DR, Conneely OM. 1992. An inducible expression system for the production of human lactoferrin in *Aspergillus nidulans*. *Gene*. 122(1):219–223.
- Yang P, Wang J, Gong G, et al. 2008. Cattle mammary bioreactor generated by a novel procedure of transgenic cloning for large-scale production of functional human lactoferrin. *PLoS ONE*. 3(10):e3453.
- Zhang J, Li L, Cai Y, Xu X, Chen J, Wu Y, Yu H, Yu G, Liu S, Zhang A, et al. 2008. Expression of active recombinant human lactoferrin in the milk of transgenic goats. *Protein Expr Purif*. 57(2):127–135.

Table S1 The composition of N-glycans form hLF after sequential exoglycosidase digestion

N-glycan Composition	neuraminidase	neuraminidase/ β-galactosidase	neuraminidase/ β-galactosidase/ β-N-acetylhexosaminidase	neuraminidase/ β-galactosidase/ β-N-acetylhexosaminidase / α-L-fucosidase
Gal ₂ GlcNAc ₂ + Man ₃ GlcNAc ₂	1663.5 Gal ₂ GlcNAc ₂ + Man ₃ GlcNAc ₂	1339.4 GlcNAc ₂ + Man ₃ GlcNAc ₂	933.3 Man ₃ GlcNAc ₂	933.3 Man ₃ GlcNAc ₂
Fuc ₁ Gal ₂ GlcNAc ₂ + Man ₃ GlcNAc ₂	1809.6 Fuc ₁ Gal ₂ GlcNAc ₂ + Man ₃ GlcNAc ₂	1485.5 Fuc ₁ GlcNAc ₂ + Man ₃ GlcNAc ₂	1079.4 Fuc ₁ + Man ₃ GlcNAc ₂	933.3 Man ₃ GlcNAc ₂
Fuc ₂ Gal ₂ GlcNAc ₂ + Man ₃ GlcNAc ₂	1955.6Fuc ₂ Gal ₂ GlcNAc ₂ + Man ₃ GlcNAc ₂	1793.6Fuc ₂ Gal ₁ GlcNAc ₂ + Man ₃ GlcNAc ₂	1590.6 Fuc ₂ Gal ₁ GlcNAc ₁ + Man ₃ GlcNAc ₂	933.3 Man ₃ GlcNAc ₂
Gal ₂ GlcNAc ₂ NeuAc ₁ + Man ₃ GlcNAc ₂	1663.5 Gal ₂ GlcNAc ₂ + Man ₃ GlcNAc ₂	1339.4 GlcNAc ₂ + Man ₃ GlcNAc ₂	933.3 Man ₃ GlcNAc ₂	933.3 Man ₃ GlcNAc ₂
Fuc ₃ Gal ₂ GlcNAc ₂ + Man ₃ GlcNAc ₂	2101.7 Fuc ₃ Gal ₂ GlcNAc ₂ + Man ₃ GlcNAc ₂	2101.7 Fuc ₃ Gal ₂ GlcNAc ₂ + Man ₃ GlcNAc ₂	2101.7 Fuc ₃ Gal ₂ GlcNAc ₂ + Man ₃ GlcNAc ₂	933.3 Man ₃ GlcNAc ₂
Fuc ₁ Gal ₂ GlcNAc ₂ NeuAc ₁ + Man ₃ GlcNAc ₂	1809.6 Fuc ₁ Gal ₂ GlcNAc ₂ + Man ₃ GlcNAc ₂	1485.5 Fuc ₁ GlcNAc ₂ + Man ₃ GlcNAc ₂	1079.4 Fuc ₁ + Man ₃ GlcNAc ₂	933.3 Man ₃ GlcNAc ₂
Fuc ₂ Gal ₂ GlcNAc ₂ NeuAc ₁ + Man ₃ GlcNAc ₂	1955.6Fuc ₂ Gal ₂ GlcNAc ₂ + Man ₃ GlcNAc ₂	1793.6Fuc ₂ Gal ₁ GlcNAc ₂ + Man ₃ GlcNAc ₂	1590.6 Fuc ₂ Gal ₁ GlcNAc ₁ + Man ₃ GlcNAc ₂	933.3 Man ₃ GlcNAc ₂
Fuc ₂ Gal ₃ GlcNAc ₃ + Man ₃ GlcNAc ₂	2320.7 Fuc ₂ Gal ₃ GlcNAc ₃ + Man ₃ GlcNAc ₂	2158.8 Fuc ₂ Gal ₂ GlcNAc ₃ + Man ₃ GlcNAc ₂	1955.7Fuc ₂ Gal ₂ GlcNAc ₂ + Man ₃ GlcNAc ₂	933.3 Man ₃ GlcNAc ₂
Fuc ₁ Gal ₂ GlcNAc ₂ NeuAc ₂ + Man ₃ GlcNAc ₂	1809.6 Fuc ₁ Gal ₂ GlcNAc ₂ + Man ₃ GlcNAc ₂	1485.5 Fuc ₁ GlcNAc ₂ + Man ₃ GlcNAc ₂	1079.4 Fuc ₁ + Man ₃ GlcNAc ₂	933.3 Man ₃ GlcNAc ₂
Fuc ₁ Gal ₃ GlcNAc ₃ NeuAc ₁ + Man ₃ GlcNAc ₂	2174.7 Fuc ₁ Gal ₃ GlcNAc ₃ + Man ₃ GlcNAc ₂	1850.7 Fuc ₁ Hex ₁ HexNAc ₃ + Man ₃ GlcNAc ₂	1444.5 Fuc ₁ Hex ₁ HexNAc ₁ + Man ₃ GlcNAc ₂	933.3 Man ₃ GlcNAc ₂
Fuc ₂ Gal ₃ GlcNAc ₃ NeuAc ₁ + Man ₃ GlcNAc ₂	2320.7 Fuc ₂ Gal ₃ GlcNAc ₃ + Man ₃ GlcNAc ₂	2158.8 Fuc ₂ Gal ₂ GlcNAc ₃ + Man ₃ GlcNAc ₂	1955.7Fuc ₂ Gal ₂ GlcNAc ₂ + Man ₃ GlcNAc ₂	933.3 Man ₃ GlcNAc ₂
Fuc ₃ Gal ₃ GlcNAc ₃ NeuAc ₁ + Man ₃ GlcNAc ₂	2466.7 Fuc ₃ Gal ₃ GlcNAc ₃ + Man ₃ GlcNAc ₂	2304.8 Fuc ₃ Gal ₂ GlcNAc ₃ + Man ₃ GlcNAc ₂	2101.8 Fuc ₃ Gal ₂ GlcNAc ₂ + Man ₃ GlcNAc ₂	933.3 Man ₃ GlcNAc ₂
Fuc ₄ Gal ₃ GlcNAc ₃ NeuAc ₁ + Man ₃ GlcNAc ₂	2612.8 Fuc ₄ Gal ₃ GlcNAc ₃ + Man ₃ GlcNAc ₂	2450.9Fuc ₄ Gal ₂ GlcNAc ₃ + Man ₃ GlcNAc ₂	2247.8Fuc ₄ Gal ₂ GlcNAc ₂ + Man ₃ GlcNAc ₂	933.3 Man ₃ GlcNAc ₂
Fuc ₁ Gal ₄ GlcNAc ₄ NeuAc ₁ + Man ₃ GlcNAc ₂	2539.8Fuc ₁ Gal ₄ GlcNAc ₄ + Man ₃ GlcNAc ₂	2053.7 Fuc ₁ Gal ₁ GlcNAc ₄ + Man ₃ GlcNAc ₂	1444.5 Fuc ₁ Hex ₁ HexNAc ₁ + Man ₃ GlcNAc ₂	933.3 Man ₃ GlcNAc ₂
Fuc ₂ Gal ₄ GlcNAc ₄ NeuAc ₁ + Man ₃ GlcNAc ₂	2685.9 Fuc ₂ Gal ₄ GlcNAc ₄ + Man ₃ GlcNAc ₂	2361.9 Fuc ₂ Gal ₂ GlcNAc ₄ + Man ₃ GlcNAc ₂	1955.7Fuc ₂ Gal ₂ GlcNAc ₂ + Man ₃ GlcNAc ₂	933.3 Man ₃ GlcNAc ₂
Fuc ₃ Gal ₄ GlcNAc ₄ NeuAc ₁ + Man ₃ GlcNAc ₂	2832.1Fuc ₃ Gal ₄ GlcNAc ₄ + Man ₃ GlcNAc ₂	2670.0 Fuc ₃ Gal ₃ GlcNAc ₄ + Man ₃ GlcNAc ₂	2466.7 Fuc ₃ Gal ₃ GlcNAc ₃ + Man ₃ GlcNAc ₂	933.3 Man ₃ GlcNAc ₂
Fuc ₄ Gal ₄ GlcNAc ₄ NeuAc ₁ + Man ₃ GlcNAc ₂	2977.9 Fuc ₄ Gal ₄ GlcNAc ₄ + Man ₃ GlcNAc ₂	2816.1 Fuc ₄ Gal ₃ GlcNAc ₄ + Man ₃ GlcNAc ₂	2612.9 Fuc ₄ Gal ₃ GlcNAc ₃ + Man ₃ GlcNAc ₂	933.3 Man ₃ GlcNAc ₂

Table S2 The composition of N-glycans form rhLF after sequential exoglycosidase digestion

N-glycan Composition	neuraminidase	neuraminidase/ β-galactosidase	neuraminidase/ β-galactosidase/ β-N-acetylhexosaminidase	neuraminidase/ β-galactosidase/ β-N-acetylhexosaminidase α-mannosidase
Hex ₂ + Man ₃ GlcNAc ₂	1257.3 Hex ₂ + Man ₃ GlcNAc ₂	1257.3 Hex ₂ + Man ₃ GlcNAc ₂	1257.3 Hex ₂ + Man ₃ GlcNAc ₂	933.5 Man ₃ GlcNAc ₂
Hex ₃ + Man ₃ GlcNAc ₂	1419.632 Hex ₃ + Man ₃ GlcNAc ₂	1419.632 Hex ₃ + Man ₃ GlcNAc ₂	1419.632 Hex ₃ + Man ₃ GlcNAc ₂	933.5 Man ₃ GlcNAc ₂
Hex ₁ HexNAc ₂ + Man ₃ GlcNAc ₂	1501.5 Hex ₁ HexNAc ₂ + Man ₃ GlcNAc ₂	1501.5 Hex ₁ HexNAc ₂ + Man ₃ GlcNAc ₂	1095.3 Hex ₁ + Man ₃ GlcNAc ₂	933.5 Man ₃ GlcNAc ₂
Hex ₄ + Man ₃ GlcNAc ₂	1581.4 Hex ₄ + Man ₃ GlcNAc ₂	1581.4 Hex ₄ + Man ₃ GlcNAc ₂	1581.4 Hex ₄ + Man ₃ GlcNAc ₂	933.5 Man ₃ GlcNAc ₂
Hex ₃ HexNAc ₁ + Man ₃ GlcNAc ₂	1622.5 Hex ₃ HexNAc ₁ + Man ₃ GlcNAc ₂	1460.4 Hex ₂ HexNAc ₁ + Man ₃ GlcNAc ₂	1257.3 Hex ₂ + Man ₃ GlcNAc ₂	933.5 Man ₃ GlcNAc ₂
Fuc ₁ Hex ₁ HexNAc ₂ + Man ₃ GlcNAc ₂	1647.5 Fuc ₁ Hex ₁ HexNAc ₂ + Man ₃ GlcNAc ₂	1647.5 Fuc ₁ Hex ₁ HexNAc ₂ + Man ₃ GlcNAc ₂	1241.3 Fuc ₁ Hex ₁ + Man ₃ GlcNAc ₂	1079.5 Fuc ₁ + Man ₃ GlcNAc ₂
Hex ₂ HexNAc ₂ + Man ₃ GlcNAc ₂	1663.5 Hex ₂ HexNAc ₂ + Man ₃ GlcNAc ₂	1663.5 Hex ₂ HexNAc ₂ + Man ₃ GlcNAc ₂	1257.3 Hex ₂ + Man ₃ GlcNAc ₂	933.5 Man ₃ GlcNAc ₂
Hex ₁ HexNAc ₃ + Man ₃ GlcNAc ₂	1704.5 Hex ₁ HexNAc ₃ + Man ₃ GlcNAc ₂	1542.5 HexNAc ₃ + Man ₃ GlcNAc ₂	933.5 Man ₃ GlcNAc ₂	933.5 Man ₃ GlcNAc ₂
Hex ₅ + Man ₃ GlcNAc ₂	1743.5 Hex ₅ + Man ₃ GlcNAc ₂	1743.5 Hex ₅ + Man ₃ GlcNAc ₂	1743.5 Hex ₅ + Man ₃ GlcNAc ₂	933.5 Man ₃ GlcNAc ₂
HexNAc ₄ + Man ₃ GlcNAc ₂	1745.6 HexNAc ₄ + Man ₃ GlcNAc ₂	1745.6 HexNAc ₄ + Man ₃ GlcNAc ₂	933.5 Man ₃ GlcNAc ₂	933.5 Man ₃ GlcNAc ₂
Fuc ₁ Hex ₂ HexNAc ₂ + Man ₃ GlcNAc ₂	1809.6 Fuc ₁ Hex ₂ HexNAc ₂ + Man ₃ GlcNAc ₂	1809.6 Fuc ₁ Hex ₂ HexNAc ₂ + Man ₃ GlcNAc ₂	1403.4 Fuc ₁ Hex ₂ + Man ₃ GlcNAc ₂	1079.5 Fuc ₁ + Man ₃ GlcNAc ₂
Fuc ₁ Hex ₁ HexNAc ₃ + Man ₃ GlcNAc ₂	1850.6 Fuc ₁ Hex ₁ HexNAc ₃ + Man ₃ GlcNAc ₂	1688.5 Fuc ₁ HexNAc ₃ + Man ₃ GlcNAc ₂	933.5 Man ₃ GlcNAc ₂	933.5 Man ₃ GlcNAc ₂
Fuc ₁ HexNAc ₄ + Man ₃ GlcNAc ₂	1891.6 Fuc ₁ HexNAc ₄ + Man ₃ GlcNAc ₂	1891.6 Fuc ₁ HexNAc ₄ + Man ₃ GlcNAc ₂	1079.5 Fuc ₁ + Man ₃ GlcNAc ₂	1079.5 Fuc ₁ + Man ₃ GlcNAc ₂
Hex ₆ + Man ₃ GlcNAc ₂	1905.6 Hex ₆ + Man ₃ GlcNAc ₂	1905.6 Hex ₆ + Man ₃ GlcNAc ₂	1905.6 Hex ₆ + Man ₃ GlcNAc ₂	933.5 Man ₃ GlcNAc ₂
HexNAc ₄ NeuAc ₁ + Man ₃ GlcNAc ₂	1745.6 HexNAc ₄ + Man ₃ GlcNAc ₂	1745.6 HexNAc ₄ + Man ₃ GlcNAc ₂	933.5 Man ₃ GlcNAc ₂	933.5 Man ₃ GlcNAc ₂
Fuc ₁ Hex ₁ HexNAc ₃ NeuAc ₁ + Man ₃ GlcNAc ₂	1850.6 Fuc ₁ Hex ₁ HexNAc ₃ + Man ₃ GlcNAc ₂	1688.5 Fuc ₁ HexNAc ₃ + Man ₃ GlcNAc ₂	933.5 Man ₃ GlcNAc ₂	933.5 Man ₃ GlcNAc ₂
Fuc ₁ HexNAc ₄ NeuAc ₁ + Man ₃ GlcNAc ₂	1891.6 Fuc ₁ HexNAc ₄ + Man ₃ GlcNAc ₂	1891.6 Fuc ₁ HexNAc ₄ + Man ₃ GlcNAc ₂	1079.5 Fuc ₁ + Man ₃ GlcNAc ₂	1079.5Fuc ₁ + Man ₃ GlcNAc ₂

Figure S1: Analysis of the sialic acid linkage using lectin blot . The “DIG Glycans Differentiation Kit” was used to identify specific carbohydrate motives. (A) SNA: recognizes α -2, 6-linked sialic acid and (D) MAA: recognizes α -2, 3-linked sialic acid. TF (Transferrin) and Fetuin were used as positive control.

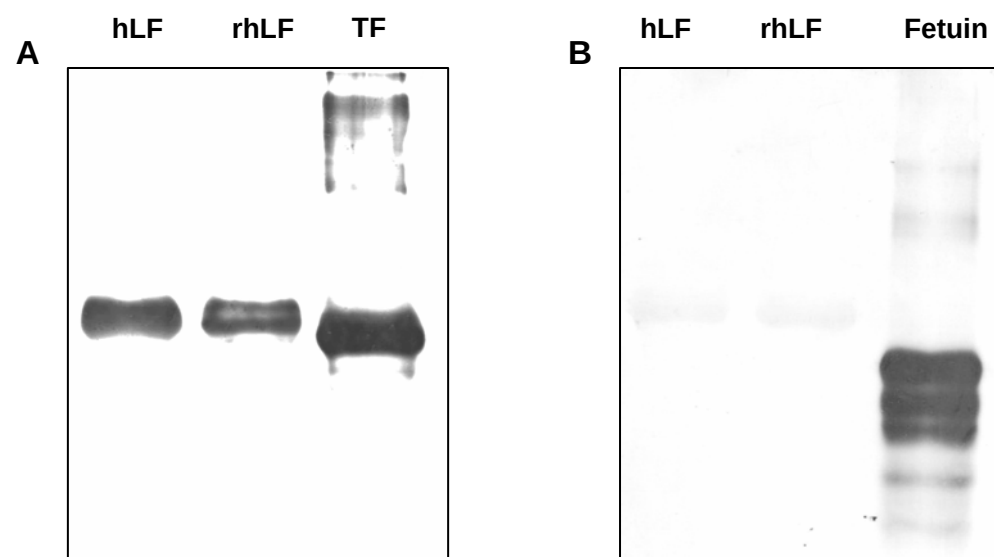


Figure S2: MALDI-TOF MS spectra of the PNGase F–released *N*-glycans from rhLF in the negative linear mode. Using the GlycoMode tools, only three acidic glycans are indentified.

