Lectin microarray profiling demonstrates equivalent global glycosylation for whey protein ingredients enriched with α-lactalbumin and milk fat globule membrane

Marie Le Berre,[1,2] Yousef Joubran,[3] Lokesh Joshi,[2] Jonathan O'Regan,[3] Michelle Kilcoyne[1,2,*]

[1] Carbohydrate Signalling Group, School of Biological and Chemical Sciences, University of Galway, Galway, Ireland

[2] Advanced Glycoscience Research Cluster, School of Biological and Chemical Sciences, University of Galway, Galway, Ireland

[3] Nestlé Development Centre Nutrition, Askeaton, Co. Limerick, Ireland

* corresponding author. Email: michelle.kilcoyne@nuigalway.ie

Abstract

Human milk fat globule membrane (MFGM) and whey proteins are nutritionally and functionally valuable, with many beneficial bioactivities associated with their glycosylation. However glycosylation of milk components other than free milk oligosaccharides are underinvestigated. Whey protein concentrate (WPC) ingredients with various enrichments or depletions are used in infant formula (IF) formulations to contribute to human milk equivalence and bioactivity benefits, but their overall or global glycosylation has not been compared. We compared the global glycosylation of commercial WPC ingredients for use in various IF formulations; two MFGM-enriched WPC ingredients (high fat HF1 and lower fat HF2), an α-lactalbumin-enriched WPC (WPC Lac) which has α-lactalbumin concentration closer to human milk and significantly less β-lactoglobulin which is not present in human milk, and two base WPC ingredients (WPC 80 and WPC 35) using lectin microarray profiling. WPC Lac and WPC HF1 glycosylation were highly similar to each other and both somewhat similar to WPC 35, while WPC HF2 was more similar to the base WPC 80 ingredient. N-linked glycosylation analysis demonstrated that WPC HF1 and WPC Lac were qualitatively most similar to one another, with WPC 80 and WPC 35 having similar structures, confirming lectin microarray profiling as a valuable method to compare global glycosylation. Thus WPC Lac may be a valuable ingredient for providing equivalent glycosylation to MFGM supplementation.

Keywords

29 Glycosylation, whey protein concentrate, α-lactalbumin, milk glycoproteins, milk fat globule
30 membrane, N-linked glycans

31

1. Introduction

Human milk is mainly comprised of water, carbohydrate, proteins, and lipids, and it is the optimal food for infant nutrition, growth, development, and health (Andreas, Kampmann, & Mehring Le-Doare, 2015). Understanding and simulating the composition of human milk in bovine milk-based infant formula (IF) is the primary focus of IF research and development. Human milk lipids contribute up to 55% of the energy of breast milk and are packaged together with glycolipids and glycoproteins in milk fat globule membranes (MFGM). MFGM proteins comprise 1-4% of the total milk protein content, and several glycoproteins which are typically found in whey fractions such as lactoferrin and α-lactalbumin (α-Lac) are found in minor abundance in MFGM (Lee et al., 2018). The carbohydrate contribution in mature human milk is from lactose, free oligosaccharides, and oligosaccharides conjugated to lipids (glycolipids) and proteins (glycoproteins) (O'Riordan, Kane, Joshi, & Hickey, 2014). Free oligosaccharide structures have been extensively studied in recent years and are responsible for many bioactivities including regulating the infant immune system, preventing pathogen adhesion, and functioning as prebiotics for the infant gut microbiome (Oliveira, Wilbey, Grandison, & Roseiro, 2015). Glycosylated components in human milk, including those from MFGM and whey glycoproteins, are also associated with various important bioactivities (Fenelon, Hickey, Buggy, McCarthy, & Murphy, 2019; O'Riordan et al., 2014; Ross et al., 2016; Takimori et al., 2011). For example, the glycosylation of lactoferrin, a major glycoprotein of whey and MFGM, has roles in iron binding and protection from infection mainly associated with sialylation (O'Riordan et al., 2014; Vega-Bautista et al., 2019). However, studies of milk glycoprotein glycosylation and bioactivities are sparse in comparison to those of human free milk oligosaccharides.

Caseins and whey proteins are the main groups of proteins in breast milk (Andreas et al., 2015). Mature human milk protein is whey protein predominant (40% casein:60% whey by approximately one month post-partum) so bovine milk proteins are adjusted to this distribution for use in IF applications as bovine milk is comprised of only 20% whey protein (Donovan, 2019). Milk glycoprotein glycosylation is altered during lactation, probably with the changing nutritional and bioactivity needs of the infant (Fenelon, Hickey, Buggy, McCarthy, & Murphy, 2019; O'Riordan et al., 2014; Ross et al., 2016; Takimori et al., 2011; Lu et al., 2019; Valk-Weeber, Eshuis-de Ruiter, Dijkhuizen, & van Leeuwen, 2020a). Lu et al. (2019) reported changes in the N- and O-linked glycan structure and composition of human milk glycoproteins

over lactation time that appeared to complement changes in the abundance of free milk oligosaccharides. Although the biological impact and mechanism for these alterations are not yet understood, glycoprotein glycosylation may provide a significant contribution to the overall nutritional value and bioactivity of milk to the infant. Given the importance of glycosylation, it may therefore be of interest to profile the overall or global glycosylation from IF ingredients.

MFGM is an important contributor to health benefits in human milk and MFGM supplementation of IF has been associated with developmental benefits including improved cognitive development and decreased infectious morbidity (Timby, Domellöf, Lönnerdal, & Hernell, 2017). Many MFGM bioactivities are associated with its glycosylation, for example providing decoy glycan ligands for pathogen receptors to prevent gastrointestinal tract infections from rotavirus and *Escherichia coli* (Manoni, Di Lorenzo, Ottoboni, Tretola, & Pinotti, 2020; Ross, Lane, Kilcoyne, Joshi, & Hickey, 2015). Nevertheless, milk fat (including MFGM) is often separated during bovine milk processing and vegetable oils are added as the source of fat in IF design. Whey protein concentrate (WPC) prepared from bovine milk is used as an ingredient for a variety of nutritional products, including IF, and is recognised for its nutritional value and beneficial bioactivities (Fenelon et al., 2019). Glycoproteins comprise the majority of whey protein and several MFGM glycoprotein and lipid components such as mucins, fatty acid synthase, butyrophilin, and phospholipids are present in WPC ingredients due to the manufacturing method (Fenelon et al., 2019; O'Riordan et al., 2014; Bansal & Bhandari, 2016; Fenelon et al., 2019; Svanborg, Johansen, Abrahamsen, & Skeie, 2015). Indeed, WPC can be depleted and/or enriched in its various components during manufacturing to produce a range of different WPC ingredients suitable for various nutritional applications. For example, in an effort to provide similar bioactivity benefits to MFGM, recently various MFGM protein- and lipid-enriched WPC ingredients have been developed for use in nutritional and IF products (Manoni et al., 2020; Timby et al., 2017). Further, WPC ingredient enriched with the glycoprotein α-lactalbumin (α-Lac) (WPC Lac), which is the most abundant protein in mature human milk, comprising 36% of human milk whey protein but only 17% of bovine milk whey protein (Barone, Moloney, O'Regan, Kelly, & O'Mahony, 2020; Donovan, 2019), is another desirable nutritional ingredient. The membrane filtration followed by selective calcium precipitation method used to enrich α-Lac also increases lipid (including phospholipid) content compared to base WPC ingredients (Barone et al., 2020), suggesting that it may additionally serve in IF design as a source of MFGM components. However the influence of

the various enrichments on the overall or global glycosylation of WPC ingredients has not been compared to date.

Providing different WPC ingredients with equivalent global glycosylation could be an important potential contributor to IF design strategies. In this study, we investigated and compared the global glycosylation of three WPC ingredients enriched in various MFGM components (two high fat MFGM-component- and one α-Lac-enriched) and two base WPC ingredients (WPC 80 and WPC 35) unaltered in their protein profiles for reference (Tables S1 and S2). MFGM components are considered desirable for IF formulations (Timby, Domellöf, Lönnerdal, & Hernell, 2017) while the two WPC base ingredients are common starting points for manufacturing various enriched WPC ingredients, or can be used as base ingredients for various food products including some IFs. The WPC Lac ingredient had been previously demonstrated to include a high lipid content (Barone et al., 2020). The WPC ingredients were first assayed for carbohydrate and protein content and their protein profiles characterised by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE). Using lectin microarray profiling, we compared the global glycosylation of the WPC ingredients for similarities between ingredient type and between ingredient batches. We also examined the N-linked glycans of the various WPC ingredients using enzymatic release and high performance liquid chromatography (HPLC) analysis to compare structural differences and similarities which contributed to global glycosylation.

2. Materials and Methods

2.1. Materials

Production batches of five commercially available WPC ingredient powders of different protein and fat content were sourced within the EU, USA and New Zealand, all spray dried powders from sweet whey (Tables S1 and S2 for macronutrient composition). Three batches (B1-B3) each of an α-Lac-enriched WPC (WPC Lac) were included in this study as well as two MFGM-enriched WPC ingredients (WPC HF1 and WPC HF2), which differed in their lipid quantity (17-20% range, and 16% average, respectively). WPC Lac was manufactured by applying membrane filtration to remove low molecular weight constituents (e.g., lactose and minerals), and then selectively precipitating α-Lac by adjustment of pH, temperature and ionic

strength (Barone et al., 2020). WPC HF1 and WPC HF2 were manufactured by membrane microfiltration to concentrate the lipid fraction. Reference WPC ingredients were included which differed in their relative protein composition (WPC 35, 35% protein, three batches B1-B3, and WPC 80 (80% protein, two batches B1 and B2, both manufactured by applying ultrafiltration to sweet whey material to concentrate the protein fraction). A quantity of each sample was fully dissolved in 3 mL phosphate buffered saline (PBS), pH 7.4, and aliquots were stored at -20 °C. One aliquot was thawed for use per experiment and not re-used.

Pure lectins were purchased from EY Laboratories, Inc. (San Mateo, CA, USA) or Vector Laboratories Ltd. (Burlingame, CA, USA). Amicon® Ultra 3 kDa molecular weight cut off (MWCO) centrifugal filters were from Merck-Millipore (Co. Cork, Ireland). Ten kDa MWCO Nanosep® centrifugal devices were from Pall Corporation, (Carrigaline, Co. Cork, Ireland). The monosaccharide 5-*N*-acetylneuraminic acid (Neu5Ac) was purchased from Dextra Laboratories Ltd. (Reading, U.K.). Alexa Fluor™ 555 NHS ester (succinimydyl ester), Pierce™ bicinchoninic acid (BCA) Protein Assay Kit, Pierce™ Silver Stain Kit, NuPAGE™ 4–12% Bis-Tris 1.0 mm 15 well precast gels and NuPAGE™ MOPS buffer were from ThermoFisher Scientific (Dublin, Ireland). Nexterion® slide H microarray slides were obtained from Schott AG (Mainz, Germany). Trypsin (cat. no. P8401), peptide-N^4-(*N*-acetyl-β-glucosaminyl)asparagine amidase F (PNGase F) (glycerol-free, cat. no. P0705), and α1-2,3,6-mannosidase (cat. no. P0768) were obtained from New England BioLabs, Inc. (Ipswich, MA, U.S.A.). GlycoClean H and S cartridges were from ProZyme (Agilent Technologies Ireland, Ltd., Co. Cork, Ireland). All other reagents were from Sigma-Aldrich Co. (Dublin, Ireland), unless otherwise noted, and were of the highest grade available.

2.2. Characterisation of WPC ingredients

Protein content of the WPC ingredients was determined using the Pierce™ BCA protein assay kit (Smith et al., 1985) using bovine serum albumin (BSA) as the standard. Neutral carbohydrate content was estimated using the Monsigny method (Monsigny, Petit, & Roche, 1988) using glucose (Glc) as the standard. Total sialic acid content was assayed using the periodate-resorcinol assay (Bhavanandan & Sheykhnazari, 1993), using Neu5Ac as the standard. All assays were carried out in triplicate in the same microtitre plate and the mean value reported.

2.3. SDS-PAGE protein profile analysis

WPC ingredients (20 μg as determined by BCA assay) were mixed with loading buffer containing 25% β-mercaptoethanol, denatured at 100 °C for 5 min and electrophoresed in NuPAGE™ 4–12% Bis-Tris gels using NuPAGE™ MOPS running buffer at 150 V constant for approximately 1 h. Gels were stained with 0.05% (w/v) Coomassie G-250 in 30% ethanol and 10% acetic acid for 1 h, partially de-stained with 15.4 MΩ (ultrapure) water overnight and scanned using a desktop scanner (CanoScan LIDE 90, Canon, Middlesex, U.K.). Protein band intensities were analysed from saved .tif image files using ImageJ Software (National Institutes of Health, Bethesda, MD, USA).

2.4. Fluorescent labelling of WPC ingredients

The WPC ingredients (0.75 mg each) were fluorescently labelled in the dark *via* amine groups using Alexa Fluor™ 555 NHS ester (λ_{ex} 555 nm, λ_{em} 572 nm; 10 μg per sample), essentially as previously described (Ross et al., 2016). The absorbance of the recovered labeled samples at 280 and 555 nm was used to calculate the relative protein concentrations and label substitution according to manufacturer's instructions using arbitrary values of M_r 100,000 and ε 100,000 M^{-1} cm^{-1} (Ross et al., 2016).

2.5. Lectin microarray construction, incubation and scanning

A panel of 50 lectins (Table S3) was printed on Nexterion® H microarray slides in six replicates per subarray and eight replicate subarrays per slide essentially as previously described (O'Riordan et al., 2014). All microarray incubations were carried out in the dark. Fluorescently labelled WPC ingredients (1 μg protein/ mL Tris-buffered saline supplemented with 1 mM each of Ca^{2+} and Mg^{2+} ions (TBS; 20 mM Tris-HCl, 100 mM NaCl, 1 mM $CaCl_2$, 1 mM $MgCl_2$, pH 7.2) with 0.05% Tween-20 (TBS-T)) were incubated on lectin microarrays essentially as previously described (O'Riordan et al., 2014; Ross et al., 2016). Titration was first carried out for each sample to determine the optimum concentration. Sample incubations (23 °C, 4 rpm, 1 h) were performed in triplicate. The microarray slides were washed twice in TBS-T for 2 min each and once with TBS. Microarrays were dried by centrifugation (450 x *g*, 5 min) and scanned immediately (Tecan PowerScanner™; 532 nm channel, full power, 90% and 40% PMT, 5 μm resolution).

2.6. Data extraction and analysis

Raw intensity values were extracted was from the image *.tif files using GenePix Pro v6.1.0.4 (Molecular Devices, Berkshire, UK) essentially as previously described (O'Riordan et al., 2014; Ross et al., 2016). Local background-corrected median feature intensity data (F532median-B532) was analysed in Excel (Version 2016, Microsoft). The median of six replicate spots per subarray was handled as a single data point for graphical and statistical analysis. Data were normalised to the per-subarray mean total intensity value of the three replicate microarray slides and binding data was presented as bar charts of average intensity of three experimental replicates +/- one standard deviation (SD). Unsupervised clustering of labelled sample binding data was performed with Hierachical Clustering Explorer v3.5 (University of Maryland, USA; http://www.cs.umd.edu/hcil/hce/). Normalised data was clustered with no pre-filtering, complete linkage, and Euclidian distance.

2.7. N-linked glycan release, labelling, and HPLC analysis

Initially proteins were denatured and precipitated from WPC ingredient solutions using a method adapted from (Yang et al., 2016) and (Cao et al., 2019). In brief, WPC ingredients were diluted in 100 mM Tris-HCl, pH 7.6, 2% SDS, and incubated on ice for 60 min with periodic sonication and vortexing. Samples were then incubated at 95 °C for 5 min, centrifuged at 12,000 *g* for 15 min and four volumes of cold acetone added to the supernatant and proteins precipitated overnight at -20 °C. After centrifugation (40 min, 14,000 *g*, 4 °C), the supernatant was discarded and the pellets air-dried. Dried pellets were then resuspended in ultrapure water and protein content quantified by BCA assay (Smith et al., 1985). N-linked glycans were then released from 40 µg precipitated protein by simultaneous digestion with trypsin and PNGase F for 3 h at 37 °C following manufacturer's instruction. Briefly, protein preparations were boiled in denaturing buffer (0.5% SDS, 40 mM DTT) for 10 min and 1% NP-40 was added to deactivate the SDS. The denatured protein preparation was then adjusted to a final buffer content of 50 mM sodium phosphate, pH 7.5 and the tryspin and PNGase F added to release the N-linked glycans. Released N-linked glycans were purified using GlycoClean H cartridges following manufacturer's instructions, evaporated to dryness, and stored at -20 °C until use. N-linked glycans were then labelled with 2-aminobenzamide (2-AB) as previously described (Bigge et al., 1995) and excess label was removed using GlycoClean S cartridges according to manufacturer's instructions. Labelled glycans were then evaporated to dryness and dissolved in 100 µL 80% acetonitrile/20% water. Ten µL of dissolved labelled glycan sample was injected on to a GlycoSep™ N-Plus column (ProZyme, Agilent Technologies; 4.6 x 150 mm,

3 μm) on a Waters 2695 HPLC instrument and eluate monitored by fluorescence detection (λ_{ex} 330 nm, λ_{em} 420 nm) using a 2475 Waters fluorescence detector. Labelled samples were separated using a linear gradient of 20-53% 50 mM ammonium formate, pH 4.4 in acetonitrile over 48 min at a flow rate of 0.67 mL/min using Waters Empower™ software. All labelled glycan samples and a reference dextran ladder were labelled and analysed simultaneously to ensure similar experimental and analysis conditions. Structures were determined from the assigned glucose unit (GU) value for each peak referencing the online GlycoStore public database (www.glycostore.org) (Zhao et al., 2018), and previously published work (Valk-Weeber, Deelman-Driessen, Dijkhuizen, Eshuis-de Ruiter, & van Leeuwen, 2020; Valk-Weeber, Eshuis-de Ruiter, Dijkhuizen, & van Leeuwen, 2020b; van Leeuwen, Schoemaker, Timmer, Kamerling, & Dijkhuizen, 2012).

Mannosidase digestion was carried out on 2-AB labelled glycans overnight at 37 °C following manufacturer's instructions. Following digestion, the mannosidase enzyme was removed by filtration using a 10 kDa MWCO Nanosep® centrifugal device according to manufacturer's instructions.

3. Results and discussion

3.1. General characterisation of WPC ingredients

WPC ingredients were initially biochemically assayed to compare their neutral sugar and sialic acid content and slight differences were noted between batches of the same ingredient (Table 1). Seasonal differences and/or lactation time are known to influence milk composition and glycosylation (Feeney et al., 2019; Parmar et al., 2020; Ross et al., 2016; Lu et al., 2019; Valk-Weeber, Eshuis-de Ruiter, Dijkhuizen, & van Leeuwen, 2020a), but these differences were typically within the standard deviation of the assays with these sample types. Neutral carbohydrate content was greatest for WPC 35 as expected compared to the other WPC ingredients (Table 1) due to the lower protein content and retention of more lactose (up to 46.5% of the dry powder) (Bansal & Bhandari, 2016). As expected, there was a similar lower neutral carbohydrate content for the other WPC ingredients, with WPC Lac demonstrating the lowest content, in agreement with the low lactose content previously determined (Barone et al., 2020). Sialic acid content can indicate the presence of the more complex type glycosylation found on glycoproteins, glycolipids and free milk oligosaccharides. WPC 35, WPC Lac and

WPC HF2 had similar quantities of sialic acid per µg protein content (Table 1), while WPC 80 and the MFGM-enriched WPC HF2 had similar content to one another and was greater than the other three ingredients.

The ratio of average neutral carbohydrate: sialic acid content could be indicative of the relative amount of complex glycosylation present. Using this ratio, WPC Lac and WPC HF1 were similar to one another (1:12 and 1:11.82, respectively), and WPC 80 and WPC HF2 were more similar to one another (1:5.4 and 1:4.81, respectively) (Table 1). However, WPC 35 apparently indicated very low relative proportion of complex glycosylation (1:0.03), so this ratio is clearly unsuitable for estimating complex glycosylation in sample as it cannot take into account potentially large contributions to the neutral carbohydrate content from lactose. On the other hand, when the ratio of average sialic acid: protein is considered, which discounts the effect of lactose, WPC 80 and WPC HF1 were more similar to one another (1:0.13 for both) and WPC 35, WPC Lac and WPC HF2 were similar (1:0.08, 1:0.07, and 1:0.08, respectively) (Table 1). This ratio might be used for comparing relative quantities of proteins with complex glycosylation in samples, but will obviously overestimate when sialic acids from free milk oligosaccharides and glycolipids are included and nor will these assays be suitable for ingredients with large quantities of lactose such as WPC 35. Critically, while these assays may be useful for comparing relative quantities of sialic acid in similar sample types, they may not be accurate for an absolute quantification as the majority of N- and O-linked glycans on mature human milk glycoproteins were reported to be non-sialylated (Lu et al., 2019). In addition, for comparing global glycosylation where structural motifs could be an important consideration, these types of colourimetric-based assays are not suitable.

3.2. Protein profiles of WPC ingredients

To compare protein profiles of WPC ingredients, the same protein quantity of each ingredient was reduced and electrophoresed on SDS-PAGE gels and stained (Fig. 1). A large number of protein bands were observed in each reduced sample (Fig. 1), confirming a wide range of proteins with a broad range of molecular weights, as expected. WPC 80 and WPC 35 had an apparently lesser number of protein bands, suggesting less diverse protein composition. However, this was due to the lack of enrichment compared to WPC Lac, WPC HF1 and WPC HF2 with their concomitant decrease in their relative proportion of β-lactoglobulin (β-Lg). In WPC 35 and WPC 80, β-Lg was the dominant protein (α-Lac:β-Lg ratio 1:3.9 and 1:5.1, respectively) as expected (Fig. S1 and Table S4) (β-Lg is band d in Fig. 1, approximate

molecular mass (Mr) 18.3 kDa; α-Lac is band e, approximate Mr 14 kDa). β-Lg is not expressed in human milk but is abundant in bovine milk. It is the main protein associated with allergies in formula-fed infants (Picariello et al., 2019), so ingredients with reduced β-Lg levels are desirable in IF applications.

The β-Lg protein apparently resolved in to two bands (approx. 17 kDa and 18 kDa) in the α-Lac- and MFGM-enriched WPC ingredient profiles in comparison to the more typical single thick band observed in the WPC 35 and WPC 80 profiles (Fig. 1). This may have been due to the reducing conditions and partial degradation of disulfide-linked aggregates of whey proteins BSA, α-Lac and β-Lg formed during ingredient processing and the high heating temperature in preparation for electrophoresis (Havea, Singh, & Creamer, 2001; Rahaman, Vasiljevic, & Ramchandran, 2015). However it is more likely that the lower overall relative quantity of β-Lg in the α-Lac- and MFGM-enriched WPC ingredients instead allowed visualisation of the approximately 17 kDa glycoprotein proteose peptone 3 (PP3, also known as glycosylation-dependent cellular adhesion molecule 1 (GlyCAM-1) or lactophorin), one of the top six most abundant proteins in the whey fraction and a component of MFGM (Fenelon et al., 2019; Valk-Weeber, Deelman-Driessen, et al., 2020), in addition to the 18 kDa β-Lg. Nevertheless, both 17 and 18 kDa bands were included for β-Lg densitometry ratio calculations for WPC HF1, WPC HF2 and WPC Lac to maintain comparability with WPC 35 and WPC 80 and consistency with previous reports (Barone et al., 2020).

While there was less relative proportion and overall quantity of β-Lg in the two MFGM-enriched WPC ingredients compared to the references WPC 35 and WPC 80, the α-Lac:β-Lg ratio was enriched in favour of β-Lg in WPC HF1 and WPC HF2 due to the concomitant substantial depletion of α-Lac in these ingredients (1:45.9 and 1:15.2, respectively). On the other hand, WPC Lac exhibited an almost equal α-Lac:β-Lg ratio of 1:1.2 (Fig. S1 and Table S4), in agreement with overall substantial β-Lg depletion and values previously reported (Barone et al., 2020). Bovine α-Lac shares a high amino acid sequence homology with human α-Lac and thus is an excellent choice for a human α-Lac substitute (Barone et al., 2020). In addition, WPC Lac increases the availability of important amino acids such as tryptophan. Tryptophan serves as a precursor to neurotransmitter serotonin essential for infant development, supports infant growth similar to breast-fed infants, and is well tolerated by the infant gastrointestinal tract (Davis, Harris, Lien, Pramuk, & Trabulsi, 2008; Lien, Davis, Euler, & Group, 2004). Use of WPC Lac ingredients in designing IFs can bring the protein profile of

313 the resulting IF closer to human milk and facilitate the design of a lower protein formula (13-
314 14 g/L) closer to that of human milk (9 g/L).

315 Casein proteins (approx. Mr 20-40 kDa) were absent in the WPC ingredient profiles as
316 expected, except in WPC 80 (band h, Fig. 1). Intense bands from 25-30 kDa in the α-Lac- and
317 MFGM-enriched WPC ingredients were likely from the abundant proteose peptone (28 kDa)
318 and immunoglobulin light chain (25 kDa) (band c in Fig. 1). Based on intensity, the proteins
319 of 'band c' were the most abundant in WPC Lac and WPC HF1, in contrast to the greatest
320 abundance of β-Lg in the other three WPC ingredients (Table S4 and Fig. S1). The intense
321 band(s) at approximately 50 kDa (band b in Fig. 1) represents the immunoglobulin heavy chain
322 and also includes lactadherin (47, 50 kDa). Another intense band at approximately 60 kDa
323 (band a) probably includes osteopontin (approx. 60 kDa), present in medium abundance in
324 whey fractions (Valk-Weeber, Deelman-Driessen, et al., 2020). Bands a, b and c were also
325 present in WPC 80 and WPC 35 but were much less intense reflecting their lower relative
326 abundance compared to the enriched WPC ingredients. BSA (66 kDa), lactoferrin, and
327 lactoperoxidase (the latter two proteins are both approx. 78 kDa) were likely components for
328 the bands above 64 kDa in the α-Lac- and MFGM-enriched WPC ingredients (Fig. 1) (Fenelon
329 et al., 2019). The intense band just above 97 kDa in all WPC ingredients was most likely
330 xanthine dehydrogenase, a MFGM protein (Svanborg et al., 2015), consistent with the greater
331 intensity of the band in the MFGM-enriched WPC ingredients, WPC HF1 and WPC HF2. The
332 fainter higher Mr bands in all WPC ingredients except for WPC 80 (Figs. 1 and S1, bands f
333 and g) likely included MFGM components fatty acid synthase and mucins (Manoni et al., 2020;
334 Svanborg et al., 2015).

335 Overall, WPC 35 and WPC 80 were more similar to one another according to their protein
336 profile and WPC HF1 and HF1 were similar to one another. While the protein profile of WPC
337 Lac was similar to WPC HF1, it had substantially less relative proportion of β-Lg.

338 **3.3. Global glycosylation comparison of WPC ingredients**

339 All batches of the WPC ingredients were fluorescently labelled *via* the amine groups of
340 constituent proteins and incubated on lectin microarrays to profile the global glycosylation
341 (Ross et al., 2016). As only protein components of the WPC ingredients are labelled by this
342 method, only glycosylation associated with proteins could have been observed and included
343 contributions from N- and O-linked glycans present, i.e. the total or global glycoprofile.

MFGM glycoproteins are embedded in the fat globule membrane, so MFGM-associated glycolipids likely additionally contributed to the MFGM-enriched WPC generated glycoprofiles (Fig. 2A) (Ross et al., 2016; Timby et al., 2017).

The glycoprofiles generated were similar across all five WPC ingredients (Fig. 2A), with differences mainly attributed by unsupervised hierarchical clustering to varying binding intensities to lectins HPA, GNA, BanLec, TJA-II, CAA, SNA-I, MAL-II, WFA, and HHA (Fig. 2B). WPC Lac and WPC HF1 displayed the greatest glycosylation similarity to one another (68% similar), followed by WPC 35 (53% similar) (Fig. 2B). As the same protein concentration of all WPC ingredients were profiled on the lectin microarray, the protein content of WPC 35 was not represented in proportion to its relative composition. Therefore while the global glycosylation of the WPC 35 glycoconjugates were qualitatively similar to those of WPC Lac and WPC HF1, quantitatively they were in fact much less. However in terms of considering inclusion of WPC 35 in a finished product, WPC ingredient quantities are usually added based on protein content which would increase the quantity of complex glycosylation present in a finished product to similar levels to that of WPC Lac or WPC HF1. The glycosylation of the other MFGM-enriched WPC ingredient WPC HF2 was more similar to the WPC 80 reference (58% similar), rather than WPC HF1 (24% similar), and together WPC 80 and WPC HF2 formed the other main cluster in the clustered heatmap (Fig. 2B).

Based on the average protein: sialic acid content ratio, WPC Lac, WPC 35, and WPC HF2 were more similar to one another, while WPC 80 and WPC HF1 were similar to each other (Table 1). This lack of agreement with the lectin microarray profiling indicates that a ratiometric comparison based on simple quantitative assays is insufficient to deduce structural similarities of complex structures. Further, protein profile similarities did not correlate with glycosylation similarities of the WPC ingredients.

The lectin microarray profiles indicated complex N- and O-linked glycosylation in all five ingredients (Fig. 2A). The lectins BanLec, Calsepa, GNA, HHA, Con A, and PSA all bind to mannose (Man) residues (Table S3), which are part of the core structure of all N-linked glycan types (but not O-linked glycans) and directly linked to the chitobiose (GlcNAc-β-(1→4)-GlcNAc) core. Intense binding to the lectins BanLec, Calsepa, GNA and HHA, comparable in intensity with binding intensities to MAL-II and RCA-I which bind to terminal residues, indicated high mannose and/or hybrid type structures in addition to complex type N-linked glycan structures. WPC Lac and WPC HF1 demonstrated higher binding intensity to BanLec,

GNA, and HHA, with WPC 35 additionally also exhibiting higher binding to GNA. High mannose type N-linked glycans are abundant on many bovine milk glycoproteins including lactoferrin, which has exclusively N-linked glycans, and high mannose and hybrid structures are expressed on lactoperoxidase (Valk-Weeber, Deelman-Driessen, et al., 2020; van Leeuwen et al., 2012). Interestingly only WPC 80 bound to the lectin NPA which indicated the presence of an α-(1→6)-linked Man motif (Table S3), possibly from a truncated hybrid structure unique to WPC 80. The binding of all ingredients to PSA additionally suggested that fucosylated moieties were present on some high mannose or hybrid N-linked structures (Table S3).

O-Linked glycans in all ingredients was indicated by binding to AIA and ACA (Fig. 2A and Table S3). The majority of O-linked glycans on WPC Lac, WPC HF2, and WPC 80 were probably sialylated as they demonstrated no or very low binding to PNA and ABL, which do not tolerate sialylation. However very low intensity binding of WPC 35 and WPC HF1 to PNA, and of WPC Lac, WPC HF1 and WPC 80 to ABL, suggested that a small proportion of the O-linked glycans in these ingredients were non-sialylated, but this proportion of non-sialylated O-linked glyans was not present in WPC HF2 which did not bind to either PNA or ABL (Fig. 2A). Some WPC ingredient components such as mucins and glycomacropeptide (GMP) are exclusively O-glycosylated, and sialylation is essential for GMP's biological activity, including promoting probiotic growth and preventing pathogenic bacteria colonisation (Córdova-Dávalos, Jiménez, & Salinas, 2019).

N-Acetylgalactosamine (GalNAc) residues in all ingredients was indicated by binding to GalNAc-specific lectins DBA, SBA, VVA, BPA, WFA, HPA, GSL-I-A4, and ACA (Fig. 2A). ACA binds to Gal-β-(1→3)-GalNAc in the core type 1 mucin structure (Wu et al., 2008), but GalNAc is also a component of the extended *N*,*N*'-diacetyllactosamine (LacdiNAc; GalNAc-β-(1→4)-GlcNAc) on N-linked glycan antennae. The LacdiNAc moiety is a relatively abundant feature of bovine α-Lac (Valk-Weeber, Deelman-Driessen, et al., 2020), lactoferrin (van Leeuwen et al., 2012), and PP3 (Inagaki et al., 2010), all constituents of whey and MFGM, and MFGM glycoproteins including butyrophilin and CD36 (Sato, Furukawa, Greenwalt, & Kobata, 1993). WFA has been used to purify LacdiNAc-containing glycans from bovine lactoferrin and bovine MFGM glycoproteins (Sato et al., 1993; van Leeuwen et al., 2012), so the higher intensity binding of WFA and HPA to WPC Lac and WPC HF1 indicated a greater relative abundance of LacdiNAc in these ingredients.

All ingredients contained very low quantities of terminal α-linked Gal residues (low intensity binding to PA-I, GSL-I-B4, and MPA) (Fig. 2A and Table S3). The Gal-α-(1→3)-Gal epitope, not produced in humans, is present on bovine milk glycoproteins including lactoferrin in minor amounts as part of several complex and hybrid type structures (van Leeuwen et al., 2012). Gal-α-(1→3)-Gal epitope-related allergy responses in humans are mainly related to previous tick exposure and subsequent red meat consumption by sensitised individuals, although a small number of bovine and goat milk-related allergies were reported in pediatric patients (Kennedy et al., 2013). However, milk glycoprotein glycosylation has been identified as bifidogenic (Garrido et al., 2012; Vega-Bautista et al., 2019), beneficial to infant microbiome development.

LEL binds to *N*-acetylglucosamine (GlcNAc) residues of the chitobiose core in N-linked glycans. The relatively low binding of all ingredients to LEL corroborated the presence of high mannose structures as well as multi-antennary complex type N-linked structures resulting in the relative inaccessibility of the chitobiose core (Fig. 2A). Low intensity binding to GlcNAc-specific lectins sWGA, DSA and STA suggested a small proportion of GlcNAc-terminated hybrid and/or complex structures, confirmed by low binding to PHA-E, which binds to biantennary N-linked glycans with or without bisecting GlcNAc (Table S3). Complex structures terminating in *N*-acetyllactosamine (LacNAc; Gal-β-(1→4)-GlcNAc) in all ingredients was indicated by intense binding to LacNAc-specific lectins RCA-I, AMA, CAA, ECA and TJA-II (Fig. 2A). LacNAc-terminating biantennary N-linked glycans are unique to bovine IgG in whey preparations, and many of these structures are sialylated and core fucosylated (Valk-Weeber, Deelman-Driessen, et al., 2020). WPC Lac and WPC HF1 differed from the other WPC ingredients by higher intensity binding to AMA and CAA.

All ingredients demonstrated moderate binding to WGA (Fig. 2A), which binds to both GlcNAc and sialic acid (Table S3). Terminal sialic acid was confirmed by high binding to MAL-II and SNA-I, indicating relatively abundant terminal α-(2,3)- and α-(2,6)-linked sialic acid, respectively, and low intensity binding to MAL-I, confirming α-(2,3)-linked sialic acid. WPC Lac and WPC HF1 differed from the other ingredients by higher binding to SNA-I (Fig. 2A), indicating a greater proportion of α-(2,6)-linked sialic acid. PP3 sialylation is exclusively α-(2,6)-linked and the majority of sialylated structures in the overall whey glycoprofile are from PP3, but α-Lac glycans are also sialylated (Valk-Weeber, Deelman-Driessen, et al., 2020). Therefore, glycans from PP3 and α-Lac may be the major contributors to the glycoprofile differences exhibited by WPC Lac and WPC HF1.

Fucose (Fuc) in all WPC ingredients was indicated by binding to TJA-II and AAL, as well as PSA. AAL binds to α-(1→6)-linked Fuc (Table S3), which modifies the chitobiose core on bovine glycoprotein N-linked glycans such as IgG and α-Lac (Valk-Weeber, Deelman-Driessen, et al., 2020). TJA-II binds to Fuc-α-(1→2)-Gal(NAc)-R and/or terminal β-linked GalNAc (Table S3). However, a lack of binding to UEA-I (which also binds to α-(1→2)-linked Fuc) indicated that TJA-II bound to LacdiNAc structures rather than α-(1→2)-linked Fuc.

Overall, O-linked and N-linked glycosylation was indicated in all WPC ingredients, with N-linked structures comprised of high mannose, hybrid, and complex type. The complex type structures had two or more antennae with terminal sialic acid (both α-(2→6)- and α-(2→3)-linked), core α-(1→6)-linked fucosylation, terminal α-linked Gal from Gal-α-(1→3)-Gal and β-linked GalNAc residues from LacdiNAc motifs, and a small proportion of terminal GlcNAc residues, in agreement with previous reports (Valk-Weeber, Deelman-Driessen, et al., 2020; van Leeuwen et al., 2012). WPC Lac and WPC HF1 demonstrated equivalent glycosylation by lectin microarray profiling and differed from the other WPC ingredients by more intense binding to lectins demonstrating greater relative quantities of α-(2→6)-sialylation (SNA-I), LacdiNAc (HPA and WFA), and LacNAc (AMA and CAA), and high mannose structures (BanLec, GNA, and HHA).

3.4. Comparison of batch to batch glycosylation of the same WPC ingredients

Seasonal and lactation variations affect milk glycosylation (O'Riordan et al., 2014; Ross et al., 2016; Takimori et al., 2011). In addition, variations in methods employed to isolate milk components including WPC can alter the composition, distribution and protein quantities in the final product, and therefore impact on overall glycosylation (Fenelon et al., 2019; Valk-Weeber, Eshuis-de Ruiter, et al., 2020b). Thus batch glycoprofiles (Table S1) were compared to assess any glycosylation variation across batches of the same WPC ingredient (Fig. S2). The glycoprofiles generated for each batch of WPC Lac, WPC HF1, WPC HF2, and WPC 80 did not cluster by individual batch, indicating that there were no global glycosylation differences between production batches. However, hierarchical clustering analysis of the WPC 35 batch glycoprofiles demonstrated individual batch clustering (Fig. S2B), indicating that there were global glycosylation differences between production batches. WPC 35 batch 2 was the least similar to the other two batches (Fig. S2B, 18% similarity), with lower lectin binding intensity to Man-binding lectins (GNA, HHA, and Banlec), GalNAc-binding lectins (WFA, DBA, VVA, and TJA-II), sialic acid-binding lectin SNA-I, LacNAc-binding lectin CAA, and Fuc-binding

lectin AAL. These data suggested that WPC 35 batches varied in relative proportion of high mannose type glycans and complex type structures bearing α-(2→6)-linked sialylation, LacdiNAc, LacNAc, and core fucosylation.

3.5. Comparison of WPC ingredient N-linked glycans

We then examined a subset of structures which contributed to global glycosylation, N-linked glycans, to compare similarities between WPC ingredients and between batches and assess correlations with the similarities revealed by lectin profiling. Initially standard enzymatic release of N-linked glycans using PNGase F was used (van Leeuwen et al., 2012) which involves protein denaturation by boiling in SDS and DTT, complexing the SDS with NP-40 as PNGase F will not function when exposed to SDS, and treating the denatured protein with PNGase F in the appropriate buffer. However poor glycan yield was obtained for WPC 35, WPC 80, and WPC HF2 (Fig. S3), which may have been due to aggregations of the complex mixture of proteins and other components present including lipids. Thus several protein pre- and co-treatments were assessed using WPC 80 as a model to increase the release of N-linked glycans from the proteins present in the ingredients. The method which allowed the optimal release of N-glycans was then applied to all ingredients to compare their N-linked glycoprofiles.

3.5.1. Optimisation of N-linked glycan release from WPC 80

Trypsin treatment prior to PNGase F digestion improves deglycosylation efficiencies, and further improvements are reported with reductive alkylation (Kita et al., 2007). For the WPC 80, protein aggregates were denatured using 2% SDS (Cao et al., 2019; Yang et al., 2016) and sonication. Proteins in the complex solution were then precipitated by either methanol/chloroform or acetone. Methanol/chloroform treatment is typically used to remove lipids from complex samples in aqueous solution while acetone precipitation is a widely used method to precipitate protein from aqueous-based mixtures. Delipidisation and protein precipitation using methanol/chloroform prior to either typical PNGase F digestion (Fig. 3, blue) or simultaneous trypsin and PNGase F digestion (Fig. 3, black) resulted in poor glycan yield. However, acetone precipitation followed by either PNGase F digestion (Fig. 3, red) or simultaneous trypsin and PNGase F digestion (Fig. 3, light blue) displayed improved N-linked glycan yields, in contrast to a previous report showing that methanol/chloroform protein recovery from milk was more efficient than acetone precipitation (Vincent et al., 2016). As the acetone precipitation followed by simultaneous trypsin and PNGase F digestion method

demonstrated the greatest N-linked glycan yield as assessed by HPLC analysis (Fig. 3, light blue), a reductive alkylation step was added post-acetone precipitation (Fig. 3, pink). Additional reductive akylation did not alter the glycan yield (the peak areas did not increase) or the composition (the chromatogram trace was not altered) compared to acetone precipitation alone (compare Fig. 3, pink and light blue), so the more convenient acetone precipitation followed by simultaneous trypsin and PNGase F digestion method was then applied to all WPC ingredients to enable comparison of their N-linked glycan content.

3.5.2. Comparison of N-linked glycans of WPC ingredients

The N-linked glycan profiles of all five WPC ingredients were compared by HPLC analysis (Fig. 4 and S4). Peaks obtained were expressed in glucose unit (GU) values (Table S5). Potential structures associated with each peak were assigned based on the GU values annotated in the GlycoStore database (Zhao et al., 2018) and in previously published datasets which have extensively structurally elucidated the N-linked glycans of whey preparations and proteins (Valk-Weeber, Deelman-Driessen, et al., 2020; Valk-Weeber, Eshuis-de Ruiter, et al., 2020b; van Leeuwen et al., 2012). N-linked glycan structures are named in the text according to Oxford annotation and all associated structures are depicted in Table S6.

A number of peaks eluted prior to approximately 14 min on the chromatographs of WPC 35, 80, and HF2 (data not shown), indicating the carry-over of free milk oligosaccharides from the ingredients despite the initial acetone precipitation step. The lactose peak at approximately 14 min was particularly prominent for WPC 35 and WPC 80, which is not surprising for WPC 35 considering that approximately half of its composition is lactose while only 3.5% (w/w) of WPC 80 is comprised of lactose (Bansal & Bhandari, 2016). For peaks indicating N-linked glycans, GU values ranging from 2.52 to 9.50, with the more complex structures having GU values above 5.4, were detected (Fig. 4B), in agreement with previous findings (van Leeuwen et al., 2012). WPC 35 and WPC 80 had peaks before GU 5.4, more abundant in WPC 35 than WPC 80, which were absent in WPC Lac, WPC HF1 and WPC HF2 (Fig. 4A). WPC 35 and WPC 80 displayed several peaks which were attributed to the N-linked glycan structures M1 (GU 2.68, Figs. 4A, S4B peak 1, and S4E peak 2, Table S6) and M2 (GU 3.42, Figs. 4A, S4B peak 5, Table S6), and also potentially free milk oligosaccharides including 3'-sialyllactosamine (GU 3.13, Figs. 4A, S4B peak 4, and S4E peak 3) and disialyllactose (DSL; GU 4.3, Figs. 4A, S4B peak 8, and S4E peak 5, Table S6). Peak 8 of WPC 35 and peak 5 of WPC 80 at GU 4.29 (Figs. 4A, S4B and S4E) indicated the presence of the structures M3 or

F(3)M2 (Table S6). However the latter structure, which contains α-(1→3)-linked Fuc, is unlikely as only α-(1→6)-linked core fucosylation was indicated by the lectin microarray data, with no binding to the lectin LTA which has specificity for α-(1→3)-linked Fuc (Table S3, Fig. 2). In addition, only α-(1→6)-linked Fuc has been reported in whey ingredients previously (van Leeuwen et al., 2012).

In the GU>5.5 range, WPC Lac, WPC HF1 and WPC HF2 had a higher number of peaks (12, 13, and 13 peaks, respectively) compared to WPC 35 and WPC 80 (10 and 11, respectively) (Fig. 4B and Table S5). Exoglycosidase digestion was carried out on the 2AB-labelled glycans using the broad specificity α1-2,3,6-mannosidase to confirm the identity of glycans contained within the various peaks (Fig. 4C) by comparing with previous work (Valk-Weeber, Deelman-Driessen, et al., 2020; Valk-Weeber, Eshuis-de Ruiter, et al., 2020b; van Leeuwen et al., 2012). The α-mannosidase used only released mannose in terminal positions. Therefore digested structures were identified as high mannose structures while the non-high mannose antenna of hybrid and complex structures were resistant to mannosidase digestion. Complete mannosidase digestion resulted in an increase in the M1 structure in WPC 35 and WPC 80, and it's appearance in WPC Lac, WPC HF1 and WPC HF2 (Fig. S4 and Table S5). WPC 35 and WPC 80 had similar complex structures that were not digested (GU 8.16, 8.50, 9.04 and 9.50, Fig. 4C and S4, Table S5, Table S6), attributed to the structures F(6)A2G(4)1S1GalNAc1S1, A2G1GalNAc1S2, F(6)A2G2Ga1S1 and F(6)A2G2Ga1S(6)1, respectively. These structures correlated with a previous reporting of a similar N-linked glycoprofile for WPC 35 and WPC 80 (Valk-Weeber, Eshuis-de Ruiter, et al., 2020b). However, the peaks at GU 8.16, 8.50, and 9.50 were partially digested in the WPC Lac, WPC HF1 and WPC HF2 chromatograms (Fig. S4, Table S6). Partial digestion of peaks indicated that the peaks contained more than a single structure, as complex structures with the same carbohydrate components but different linkages or similar size glycans can co-elute. For example M9 from lactoferrin and the monosialylated F(6)A2G2AS(6)1 structure from IgG co-eluted at GU 9.48 (Valk-Weeber, Deelman-Driessen, et al., 2020).

Minor differences were noted between the WPC Lac, WPC HF1 and WPC HF2 chromatograms following mannosidase digestion (Fig. S4). WPC Lac and WPC HF2 had a peak at GU 9.01 which remained undigested, whereas this peak was partially digested in WPC HF1 (Fig. 4B and 4C, Table S5, Table S6), indicating the potential presence of M8 and F(6)A2G2Ga1S1, among others. WPC Lac also had one peak at GU 7.85 that was not digested whereas it was

digested for WPC HF1 and WPC HF2 (structures M7 and F(6)A2G1GalNAc1S1). On the other hand, WPC HF2 had a peak at GU 7.07 that was not mannosidase-digested (structure A2G1S1) whereas it was partially digested for WPC Lac and WPC HF1 (structure M6). The different WPC ingredients contained different proportions of high mannose structures, which are mostly present on lactoferrin (Valk-Weeber, Eshuis-de Ruiter, Dijkhuizen, & van Leeuwen, 2020a).

Comparison of the N-linked glycosylation profiles of various WPC ingredient batches demonstrated no differences in N-linked glycans between individual batches of WPC Lac, WPC HF1, WPC HF2, and WPC 80 (Fig. S5A, C, D, and E), in agreement with the conclusions from the lectin microarray data (Fig. S2). Only two batches of WPC 35 were compared for N-linked glycosylation, batches 2 and 3 (Fig. S5B), which were indicated as less similar to one another by lectin microarray profiling (Fig. S2). The relative distribution of the mannose-containing M1 structure differed between batches, in agreement with lectin binding data, but the relative distribution of more complex structures appeared to be similar between batches. These data suggest that the WPC 35 batch differences in GalNAc, sialic acid, and Fuc-containing structures indicated by lectin microarray profiling may have been due to differences in O-linked glycosylation.

Overall we suggest over 50 N-linked glycan structures were present in the various WPC ingredients (Tables S5 and S6), in agreement with previous work which described over 60 structures in WPC-based infant formula Deminal 90 (van Leeuwen et al., 2012). While the N-linked glycan chromatographs had a majority of common peaks for all five WPC ingredients, WPC 35 and WPC 80 were more similar to one another than WPC HF1, WPC Lac and WPC HF2. Of the latter three WPC ingredients, the N-linked glycoprofile of WPC HF1 and WPC Lac were more similar to one another than to WPC HF2 in terms of structures and their relative proportions (Fig. 4C). In whey protein derived preparations, the majority of the identified structures are carried by the five main glycoproteins of whey: lactoferrin, IgG, PP3, α-Lac, and lactoperoxidase, with PP3 being the major contributor (Valk-Weeber, Deelman-Driessen, et al., 2020). Other glycoproteins from added ingredients such as MFGM components or α-Lac also contribute to the overall N-linked glycosylation profile of the WPC ingredients. As N-linked glycans present on α-Lac are very similar to those present on PP3, it is therefore difficult to conclusively state which glycoproteins are responsible for the subtle differences noted between WPC Lac and WPC HF1.

4. Conclusions

Milk glycoprotein glycosylation is an important contributor to nutrition and plays important roles in the specific bioactivities of MFGM and whey glycoproteins including lactoferrin, GMP, and immunoglobulin. Despite its importance, the global glycosylation of supplemented WPC ingredients have not been previously compared. In this study, two MFGM-component- and one α-Lac-enriched WPC ingredients (containing elevated phospholipid levels) were investigated and compared by lectin microarray profiling, using unadjusted WPC 80 and WPC 35 as references.

Surprisingly the global glycosylation of WPC HF1 and HF2 were not equivalent, indicating that these ingredients are not interchangeable in terms of achieving the same glycosylation outcome in a final formulation. However, the global glycosylation of WPC Lac and WPC HF1 exhibited high similarity while the other slightly lower fat MFGM-component enriched WPC HF2 was more similar to the base WPC 80 ingredient. Detailed analysis of the WPC ingredient N-linked glycans by enzymatic release and HPLC structural analysis showed that the N-linked glycosylation of WPC Lac, WPC HF1, and WPC HF2 were similar. WPC Lac and WPC HF1 were almost qualitatively equivalent with differences most likely attributable to the glycans of α-Lac. The reference ingredients WPC 35 and WPC 80 demonstrated more similar N-linked glycan structures to one another both before and after mannosidase digestion. Based on this study, lectin microarray profiling was demonstrated to be an effective and convenient method for comparing global glycosylation similarities across samples.

Enriched WPC ingredients differed from the base WPC ingredients, demonstrating that altering protein and lipid composition can change global glycosylation. Moreover, protein profile comparisons demonstrated the similarity of WPC 35 and WPC 80 to one another, and WPC HF1 to WPC HF2. However, as lectin microarray profiling and N-linked glycan profiling confirmed that the WPC 80 glycosylation was more similar to WPC HF2 than WPC 35, these observations confirm that protein profiling did not correlate with glycosylation similarity for WPC ingredients.

Based on these data, we suggest that not all MFGM-enriched WPC ingredients available can achieve the same global glycosylation and should not be interchangeably used. On the other hand, WPC Lac demonstrated the equivalent global and N-linked glycosylation as an MFGM-

enriched WPC ingredient. As WPC Lac also has the advantage of a significantly better α-Lac:β-Lg ratio, it could facilitate the design of IFs closer to human milk.

CRediT authorship contribution statement

Marie Le Berre – conceptualisation, data curation, investigation, formal analysis, methodology, writing - original draft writing, writing - review and editing; Yousef Joubran - conceptualisation, formal analysis, funding acquisition, project administration, resources, writing - review and editing; Lokesh Joshi - conceptualisation, funding acquisition, writing - review and editing; Jonathan O'Regan - conceptualisation, formal analysis, funding acquisition, project administration, resources, writing - review and editing; Michelle Kilcoyne - conceptualisation, data curation, formal analysis, funding acquisition, methodology, project administration, supervision, writing - original draft writing, writing - review and editing.

Declaration of Competing Interest

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

Acknowledgements

The authors acknowledge Nestlé Ireland for funding this study.

References

Andreas, N. J., Kampmann, B., & Mehring Le-Doare, K. (2015). Human breast milk: A review on its composition and bioactivity. *Early Human Development, 91*(11), 629-635. https://doi.org/10.1016/j.earlhumdev.2015.08.013

Bansal, N., & Bhandari, B. (2016). Functional milk proteins: Production and utilization - whey-based ingredients. In P. L. H. McSweeney & J. A. O'Mahony (Eds.), *Advanced Dairy Chemistry* (4th ed., Vol. 1B, pp. 67-98). New York: Springer.

Barone, G., Moloney, C., O'Regan, J., Kelly, A. L., & O'Mahony, J. A. (2020). Chemical composition, protein profile and physicochemical properties of whey protein concentrate ingredients enriched in α-lactalbumin. *Journal of Food Composition and Analysis, 92*, 103546. https://doi.org/10.1016/j.jfca.2020.103546

Bhavanandan, V. P., & Sheykhnazari, M. (1993). Adaptation of the periodate-resorcinol method for determination of sialic acids to a microassay using microtiter plate reader. *Analytical Biochemistry, 213*(2), 438-440. https://doi.org/10.1006/abio.1993.1445

Bigge, J. C., Patel, T. P., Bruce, J. A., Goulding, P. N., Charles, S. M., & Parekh, R. B. (1995). Nonselective and efficient fluorescent labeling of glycans Using 2-amino benzamide and anthranilic acid. *Analytical Biochemistry, 230*(2), 229-238. https://doi.org/10.1006/abio.1995.1468

Cao, X., Zheng, Y., Wu, S., Yang, N., Wu, J., Liu, B., . . . Yue, X. (2019). Characterization and comparison of milk fat globule membrane N-glycoproteomes from human and bovine colostrum and mature milk. *Food & Function, 10*(8), 5046-5058. https://doi.org/10.1039/C9FO00686A

Córdova-Dávalos, L. E., Jiménez, M., & Salinas, E.. (2019). Glycomacropeptide bioactivity and health: A review highlighting action mechanisms and signaling pathways. *Nutrients, 11*(3), 598. https://doi.org/10.3390/nu11030598

Davis, A. M., Harris, B. J., Lien, E. L., Pramuk, K., & Trabulsi, J. (2008). α-Lactalbumin-rich infant formula fed to healthy term infants in a multicenter study: plasma essential amino acids and gastrointestinal tolerance. *European Journal of Clinical Nutrition, 62*(11), 1294-1301. https://doi.org/10.1038/sj.ejcn.1602848

Donovan, S. M. (2019). Human milk proteins: Composition and physiological significance. In S. M. Donovan, J. B. German, B. Lönnerdal & A. Lucas (Eds.), Human milk: Composition, clinical benefits and future opportunitiesNestlé Nutr Inst Workshop Ser (Vol. 90, pp. 93-101): Nestlé Nutrition Institute, Switzerland/S. Karger AG., Basel. https://doi.org/10.1159/000490298

Feeney, S., Gerlach, J. Q., Slattery, H., Kilcoyne, M., Hickey, R. M., & Joshi, L. (2019). Lectin microarray profiling and monosaccharide analysis of bovine milk immunoglobulin G oligosaccharides during the first 10 days of lactation. *Food Science & Nutrition, 7*(5), 1564-1572. https://doi.org/10.1002/fsn3.950

Fenelon, M. A., Hickey, R. M., Buggy, A., McCarthy, N., & Murphy, E. G. (2019). Whey proteins in infant formula. In H. C. Deeth & N. Bansal (Eds.), *Whey proteins: From milk to medicine* (pp. 439-494). London: Academic Press.

Garrido, D., Nwosu, C., Ruiz-Moyano, S., Aldredge, D., German, J. B., Lebrilla, C. B., & Mills, D. A. (2012). Endo-β-*N*-acetylglucosaminidases from infant gut-associated Bifidobacteria release complex *N*-glycans from human milk glycoproteins. *Molecular & Cellular Proteomics, 11*(9), 775-785. https://doi.org/10.1074/mcp.M112.018119

Havea, P., Singh, H., & Creamer, L. K. (2001). Characterization of heat-induced aggregates of β-lactoglobulin, α-lactalbumin and bovine serum albumin in a whey protein concentrate environment. *Journal of Dairy Research, 68*(3), 483-497. https://doi.org/10.1017/S0022029901004964

Inagaki, M., Nakaya, S., Nohara, D., Yabe, T., Kanamaru, Y., & Suzuki, T.. (2010). The multiplicity of N-glycan structures of bovine milk 18 kDa lactophorin (milk GlyCAM-1). *Bioscience, Biotechnology, and Biochemistry, 74*(2), 447-450. https://doi.org/10.1271/bbb.90887

Kennedy, J. L., Stallings, A. P., Platts-Mills, T. A. E., Oliveira, W. M., Workman, L., James, H. R., . . . Commins, S. P. (2013). Galactose-α-1,3-galactose and delayed anaphylaxis, angioedema, and urticaria in children. *Pediatrics, 131*(5), e1545-e1552. https://doi.org/10.1542/peds.2012-2585

Kita, Y., Miura, Y., Furukawa, J., Nakano, M., Shinohara, Y., Ohno, M., . . . Nishimura, S.-I. (2007). Quantitative glycomics of human whole serum glycoproteins based on the standardized protocol for liberating N-glycans. *Molecular & Cellular Proteomics, 6*(8), 1437-1445. https://doi.org/10.1074/mcp.T600063-MCP200

Lee, H., Padhi, E., Hasegawa, Y., Larke, J., Parenti, M., Wang, A., . . . Slupsky, C. (2018). Compositional dynamics of the milk fat globule and its role in infant development. *Frontiers in Pediatrics*, *6*, 313. https://doi.org/10.3389/fped.2018.00313Lien, E. L., Davis, A. M., Euler, A. R., & Group, Multicenter Study. (2004). Growth and safety in term infants fed reduced-protein gormula with added bovine alpha-lactalbumin. *Journal of Pediatric Gastroenterology and Nutrition, 38*(2), 170-176. https://doi.org/10.1097/00005176-200402000-00013

Lu, Y., Liu, J., Jia, Y., Yang, Y., Chen, Q., Sun, L., . . . Wang, W. (2019) Mass spectrometry analysis of changes in human milk N/O-glycopatterns at different lactation stages. *Journal of Agricultural and Food Chemistry*, *67*, 10701-10712. https://doi.org/10.1021/acs.jafc.9b02034

Manoni, M., Di Lorenzo, C., Ottoboni, M., Tretola, M., & Pinotti, L. (2020). Comparative proteomics of milk fat globule membrane (MFGM) proteome across species and lactation stages and the potentials of MFGM fractions in infant formula preparation. *Foods, 9*(9), 1251. https://doi.org/10.3390/foods9091251

Monsigny, M., Petit, C., & Roche, A.-C. (1988). Colorimetric determination of neutral sugars by a resorcinol sulfuric acid micromethod. *Analytical Biochemistry, 175*(2), 525-530. https://doi.org/10.1016/0003-2697(88)90578-7

Oliveira, D.L., Wilbey, R.A., Grandison, A.S., Roseiro, L.B. (2015). Milk oligosaccharides: A review. *International Journal of Dairy Technology*, *68*(3), 305-321. https://doi.org/10.1111/1471-0307.12209

O'Riordan, N., Kane, M., Joshi, L., & Hickey, R. M. (2014). Structural and functional characteristics of bovine milk protein glycosylation. *Glycobiology, 24*(3), 220-236. https://doi.org/10.1093/glycob/cwt162

O'Riordan, N., Gerlach, J. Q., Kilcoyne, M., O'Callaghan, J., Kane, M., Hickey, R. M., & Joshi, L. (2014). Profiling temporal changes in bovine milk lactoferrin glycosylation using lectin microarrays. *Food Chemistry, 165*, 388-396. https://doi.org/10.1016/j.foodchem.2014.05.086

Parmar, P., Lopez-Villalobos, N., Tobin, J. T., Murphy, E., McDonagh, A., Crowley, S. V., . . . Shalloo, L. (2020). The effect of compositional changes due to seasonal variation on milk density and the determination of season-based density conversion factors for use in the dairy industry. *Foods, 9*(8), 1004. https://doi.org/10.3390/foods9081004

Picariello, G., De Cicco, M., Nocerino, R., Paparo, L., Mamone, G., Addeo, F., & Berni Canani, R. (2019). Excretion of dietary cow's milk derived peptides into breast milk. *Frontiers in Nutrition, 6*(25). https://doi.org/10.3389/fnut.2019.00025

Rahaman, T., Vasiljevic, T., & Ramchandran, L. (2015). Conformational changes of β-lactoglobulin induced by shear, heat, and pH-Effects on antigenicity. *Journal of Dairy Science, 98*(7), 4255-4265. https://doi.org/10.3168/jds.2014-9010

Ross, S. A., Gerlach, J. Q., Gill, S. K., Lane, J. A., Kilcoyne, M., Hickey, R. M., & Joshi, L. (2016). Temporal alterations in the bovine buttermilk glycome from parturition to milk maturation. *Food Chemistry, 211*, 329-338. https://doi.org/10.1016/j.foodchem.2016.05.027

Ross, S. A., Lane, J. A., Kilcoyne, M., Joshi, L., & Hickey, R. M. (2015). The milk fat globule membrane *Biotechnology of Bioactive Compounds* (pp. 631-668).

Sato, T., Furukawa, K., Greenwalt, D. E., & Kobata, A. (1993). Most bovine milk fat globule membrane glycoproteins contain asparagine-linked sugar chains with GalNAcβ1→4GlcNAc groups. *The Journal of Biochemistry, 114*(6), 890-900. https://doi.org/10.1093/oxfordjournals.jbchem.a124273

Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., . . . Klenk, D. C. (1985). Measurement of protein using bicinchoninic acid. *Analytical Biochemistry, 150*(1), 76-85. https://doi.org/10.1016/0003-2697(85)90442-7

Svanborg, S., Johansen, A.-G., Abrahamsen, R. K., & Skeie, S. B. (2015). The composition and functional properties of whey protein concentrates produced from buttermilk are comparable with those of whey protein concentrates produced from skimmed milk. *Journal of Dairy Science, 98*(9), 5829-5840. https://doi.org/10.3168/jds.2014-9039

Takimori, S., Shimaoka, H., Furukawa, J.-I., Yamashita, T., Amano, M., Fujitani, N., . . . Nishimura, S.-I. (2011). Alteration of the N-glycome of bovine milk glycoproteins during early lactation. *The FEBS Journal, 278*(19), 3769-3781. https://doi.org/10.1111/j.1742-4658.2011.08299.x

Timby, N., Domellöf, M., Lönnerdal, B., & Hernell, O. (2017). Supplementation of infant formula with bovine milk fat globule membranes. *Advances in Nutrition, 8*(2), 351-355. https://doi.org/10.3945/an.116.014142

Valk-Weeber, R. L., Deelman-Driessen, C., Dijkhuizen, L., Eshuis-de Ruiter, T., & van Leeuwen, S. S. (2020). In depth analysis of the contribution of specific glycoproteins to the overall bovine whey N-linked glycoprofile. *Journal of Agricultural and Food Chemistry, 68*(24), 6544-6553. https://doi.org/10.1021/acs.jafc.0c00959

Valk-Weeber, R. L., Eshuis-de Ruiter, T., Dijkhuizen, L., & van Leeuwen, S. S. (2020a). Dynamic temporal variations in bovine lactoferrin glycan structures. *Journal of Agricultural and Food Chemistry, 68*(2), 549-560. https://doi.org/10.1021/acs.jafc.9b06762

Valk-Weeber, R. L., Eshuis-de Ruiter, T., Dijkhuizen, L., & van Leeuwen, S. S. (2020b). Quantitative analysis of bovine whey glycoproteins using the overall N-linked whey glycoprofile. *International Dairy Journal, 110*, 104814. https://doi.org/10.1016/j.idairyj.2020.104814

van Leeuwen, S. S., Schoemaker, R. J. W., Timmer, C. J. A. M., Kamerling, J. P., & Dijkhuizen, L. (2012). Use of *Wisteria floribunda* agglutinin affinity chromatography in the structural analysis of the bovine lactoferrin N-linked glycosylation. *Biochimica et Biophysica Acta (BBA) - General Subjects, 1820*(9), 1444-1455. https://doi.org/10.1016/j.bbagen.2011.12.014

Vega-Bautista, A., de la Garza, M., Carrero, J. C., Campos-Rodríguez, R., Godínez-Victoria, M., & Drago-Serrano, M. E. (2019). The impact of lactoferrin on the growth of intestinal inhabitant bacteria. *International Journal of Molecular Sciences, 20*(19), 4707. https://doi.org/10.3390/ijms20194707

Vincent, D., Ezernieks, V., Elkins, A., Nguyen, N., Moate, P. J., Cocks, B. G., & Rochfort, S. (2016). Milk bottom-up proteomics: Method optimization. *Frontiers in Genetics, 6*(360). https://doi.org/10.3389/fgene.2015.00360

Wu, A. M., Wu, J. H., Yang, Z., Singh, T., Goldstein, I. J., & Sharon, N. (2008). Differential contributions of recognition factors of two plant lectins – Amaranthus caudatus lectin and Arachis hypogea agglutinin, reacting with Thomsen-Friedenreich disaccharide (Galβ1–3GalNAcα1–Ser/Thr). *Biochimie, 90*(11), 1769-1780. https://doi.org/10.1016/j.biochi.2008.08.001

Yang, Y., Zheng, N., Wang, W., Zhao, X., Zhang, Y., Han, R., . . . Wang, J. (2016). N-glycosylation proteomic characterization and cross-species comparison of milk fat globule membrane proteins from mammals. *PROTEOMICS, 16*(21), 2792-2800. https://doi.org/10.1002/pmic.201500361

Zhao, S., Walsh, I., Abrahams, J. L, Royle, L., Nguyen-Khuong, T., Spencer, D., . . . Campbell, M. P. (2018). GlycoStore: a database of retention properties for glycan analysis. *Bioinformatics, 34*(18), 3231-3232. https://doi.org/10.1093/bioinformatics/bty319

Legends to figures

Fig. 1. 4-12% SDS-PAGE profile of the five WPC ingredients (20 μg each) under reducing conditions and stained with Coomassie G-250. Lane L: molecular weight protein marker ladder (SeeBlue^{TM} Plus2 pre-stained protein standard); Lane 1: WPC Lac B1; Lane 2: WPC Lac B2; Lane 3: WPC Lac B3; Lane 4: WPC 35 B1; Lane 5: WPC 35 B2; Lane 6: WPC 35 B3; Lane 7: WPC HF1 B1; Lane 8: WPC HF1 B2; Lane 9: WPC HF1 B3; Lane 10: WPC HF2 B1; Lane 11: WPC HF2 B2; Lane 12: WPC HF B3; Lane 13: WPC 80 B1; Lane 14: WPC 80 B2. Arrows accompanied by lower case letters a-h indicate protein bands of interest for densitometry analysis, where d indicates β-Lg and e indicates α-Lac.

Fig. 2. (A) Lectin microarray binding profiles of the five WPC ingredients. Data represents the average of three technical replicate experiments for each sample and error bars are +/- one SD. (B) Unsupervised clustering of lectin binding intensity data of all batches of the five WPC ingredients. Normalised data were subjected to unsupervised, Euclidean distance, with complete linkage clustering. Percentage similarity is indicated at nodes.

Fig. 3. Optimisation of N-glycan release from WPC 80. (A) Chromatograms of the five different release conditions assessed from approximately 10-45 min. Blue, methanol/chloroform delipidisation followed by PNGase F digestion; black, methanol/chloroform delipidisation followed by simultaneous trypsin/PNGase F digestion; red, acetone protein precipitation followed by PNGase F digestion; light blue, acetone protein precipitation followed by simultaneous trypsin/PNGase F digestion; pink, acetone protein precipitation followed by reductive alkylation and PNGase F digestion. (B) Zoom-in of the peaks representing the N-linked glycan structures in the chromatograms.

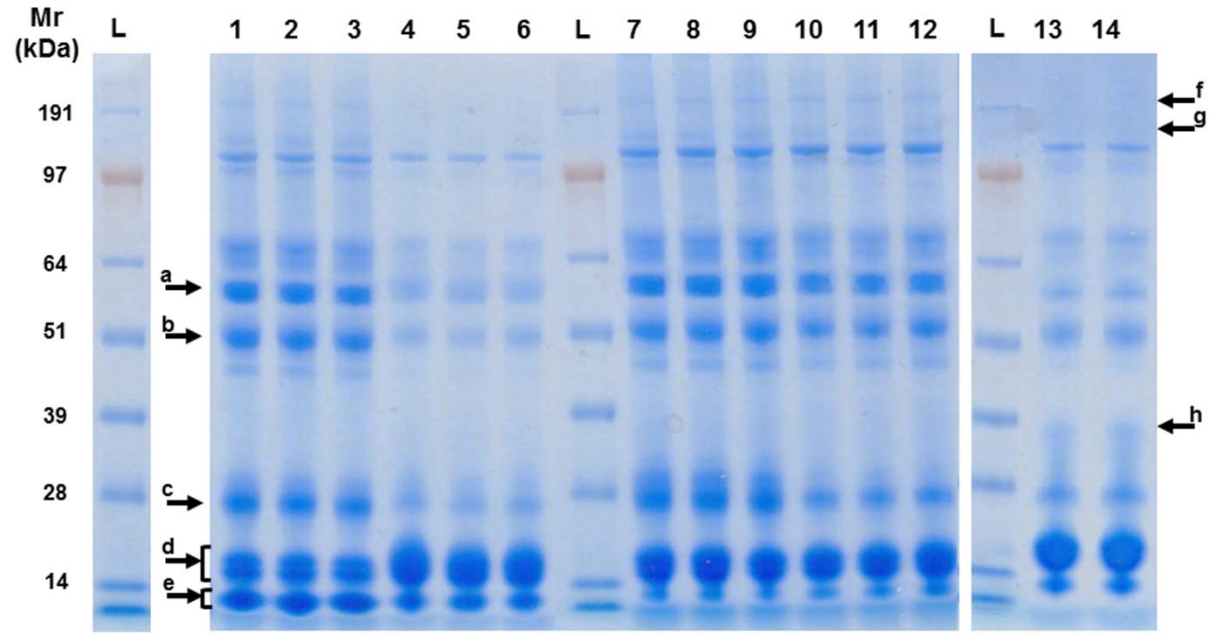
Fig. 4. NP-HPLC chromatograms of 2-AB-labelled N-linked glycans released from WPC 35 (blue), WPC 80 (pink), WPC Lac (black), WPC HF1 (red), WPC HF2 (light blue). GU scale is based on the elution of the 2-AB labelled dextran ladder. (A) Chromatograms from approximately 14-45 min. (B) Zoom-in of N-linked glycan peaks in the chromatograms. (C) Zoom-in of N-linked glycan peaks on the chromatograms after mannosidase digestion.

835 **Table 1.** Neutral carbohydrate and sialic acid content of the WPC ingredient powders analysed
836 (in μg/μg protein content of the mean of three technical replicates estimated by BCA assay with
837 +/- 1 standard deviation of the technical triplicate measurement in parenthesis), and ratio of
838 average sialic acid content to neutral carbohydrate content and protein content. Av, average.

Sample code	Batch	Neutral carbohydrates	Sialic acid	Ratio carb:sialic acid	Ratio protein:sialic acid
WPC 35	B1	0.342 (0.143)	0.081 (0.008)		
	B2	0.432 (0.119)	0.072 (0.016)		
	B3	0.320 (0.018)	0.093 (0.004)		
	Av	0.365 (0.063)	0.082 (0.010)	1:0.23	1:0.08
WPC 80	B1	0.024 (0.005)	0.158 (0.038)		
	B2	0.028 (0.004)	0.092 (0.003)		
	Av	0.026 (0.003)	0.125 (0.047)	1:4.81	1:0.13
WPC Lac	B1	0.005 (0.000)	0.072 (0.007)		
	B2	0.007 (0.001)	0.067 (0.012)		
	B3	0.007 (0.000)	0.078 (0.006)		
	Av	0.006 (0.001)	0.072 (0.006)	1:12.00	1:0.07
WPC HF1	B1	0.012 (0.001)	0.142 (0.002)		
	B2	0.011 (0.004)	0.120 (0.013)		
	B3	0.009 (0.003)	0.128 (0.023)		
	Av	0.011 (0.002)	0.130 (0.011)	1:11.82	1:0.13
WPC HF2	B1	0.018 (0.000)	0.072 (0.008)		
	B2	0.014 (0.002)	0.081 (0.003)		
	B3	0.013 (0.000)	0.090 (0.012)		
	Av	0.015 (0.003)	0.081 (0.009)	1:5.40	1:0.08

839

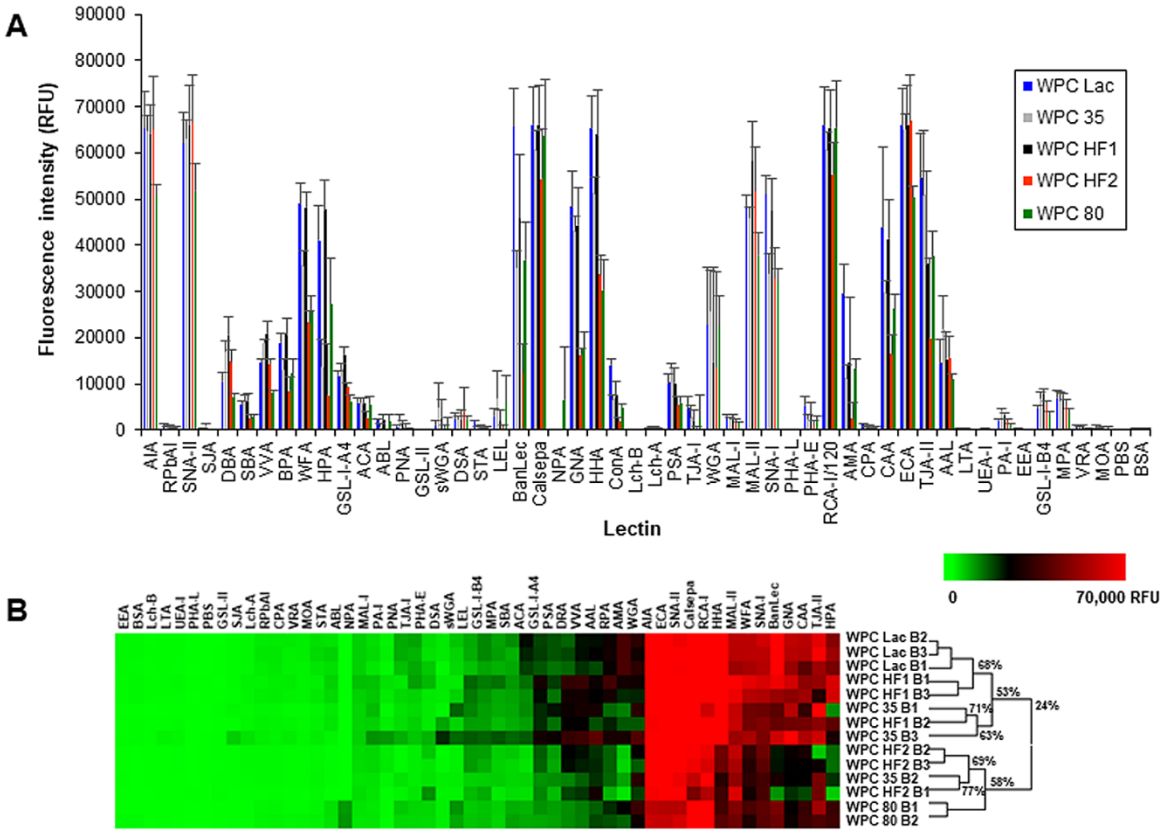
840



841

842 Figure 1

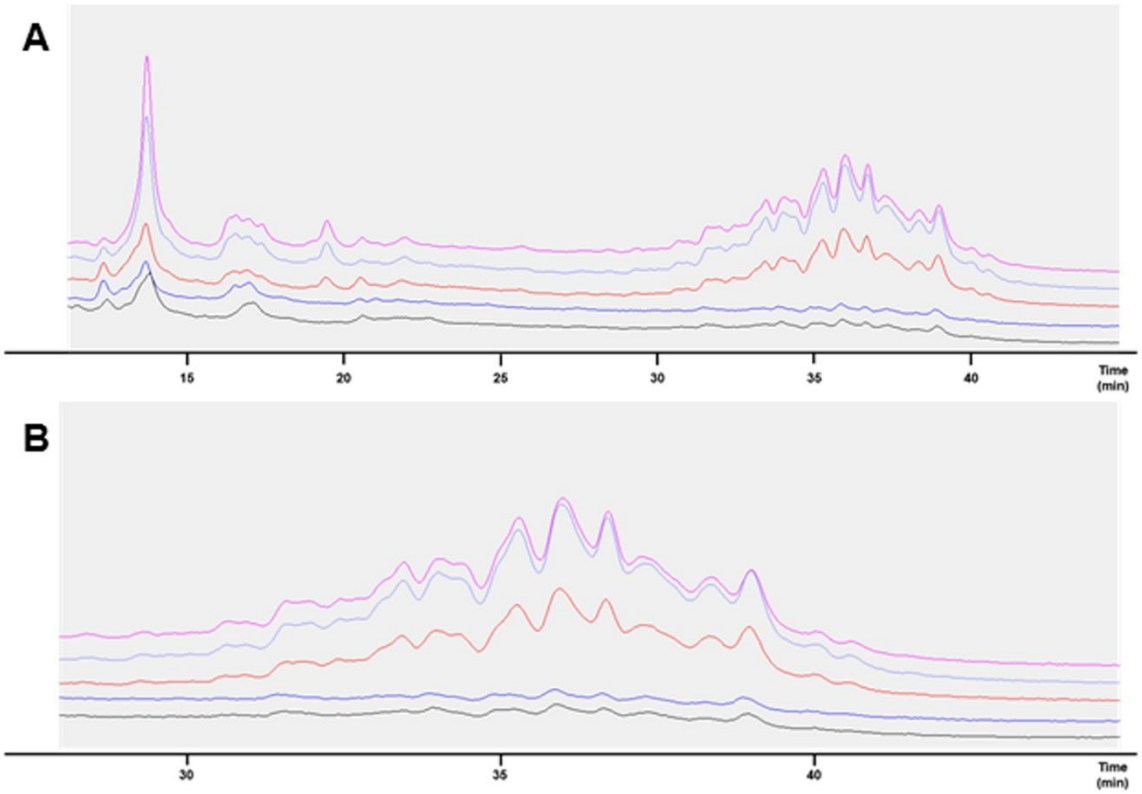
843



844

845 Figure 2

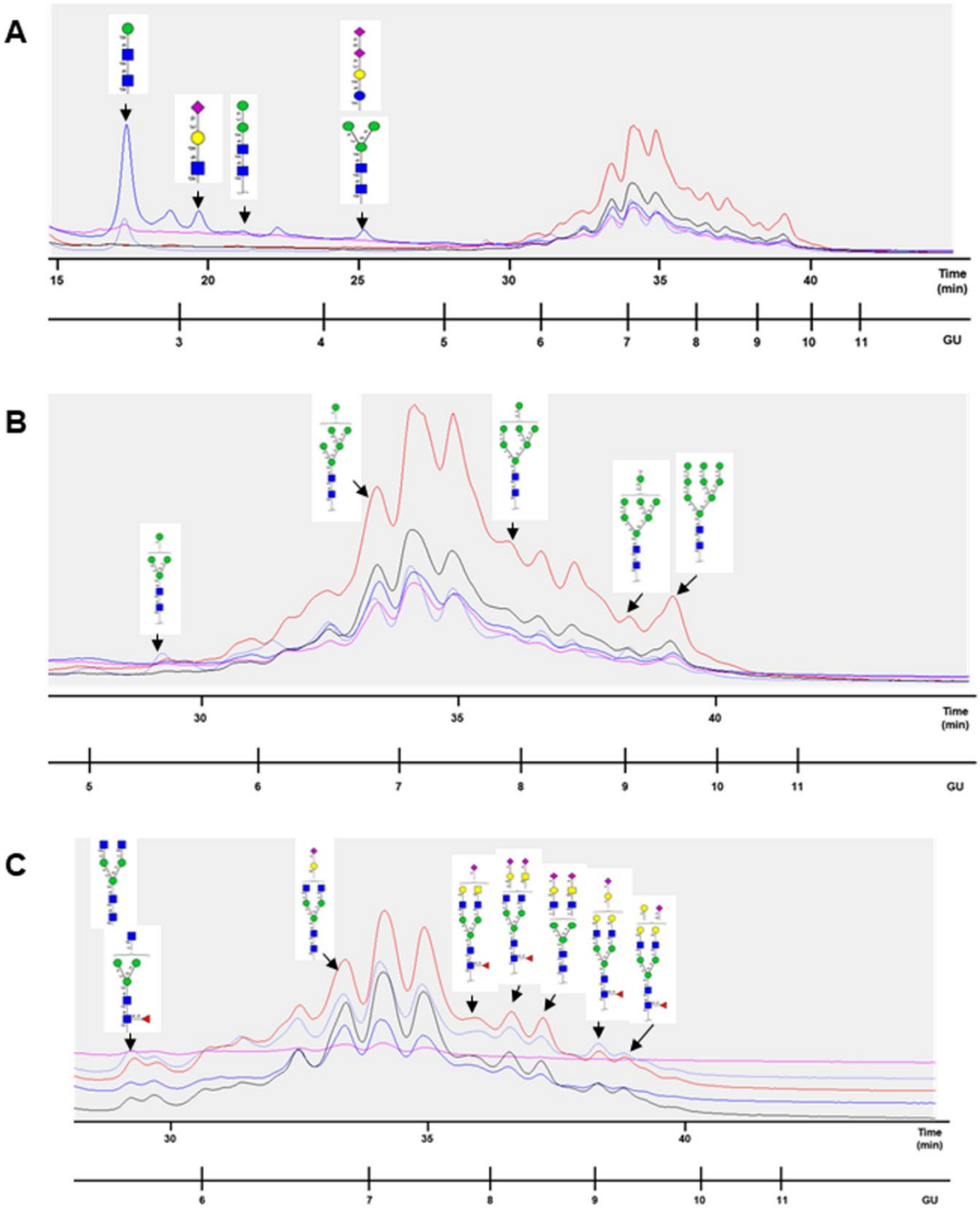
846



847

848 Figure 3

849



850

851 Figure 4

Supplementary information

Lectin microarray profiling demonstrates equivalent global glycosylation for whey protein ingredients enriched with α-lactalbumin and milk fat globule membrane

Marie Le Berre,[1,2] Yousef Joubran,[3] Lokesh Joshi,[2] Jonathan O'Regan,[3] Michelle Kilcoyne[1,2]

[1] Carbohydrate Signalling Group, School of Biological and Chemical Sciences, University of Galway, Galway, Ireland

[2] Advanced Glycoscience Research Cluster, School of Biological and Chemical Sciences, University of Galway, Galway, Ireland

[3] Nestlé Development Centre Nutrition, Wyeth Nutritionals Ireland, Askeaton, Co. Limerick, Ireland

Table of contents

Table S4. Integrated densitometry analysis of Coomassie-stained SDS-PAGE protein profiles of the five WPC ingredients.

Table S5. GU values and assigned potential N-linked glycan structures from the five WPC ingredients.

Table S6. Predicted structures present in WPC 35, WPC 80, WPC Lac, WPC HF1 and WPC HF2.

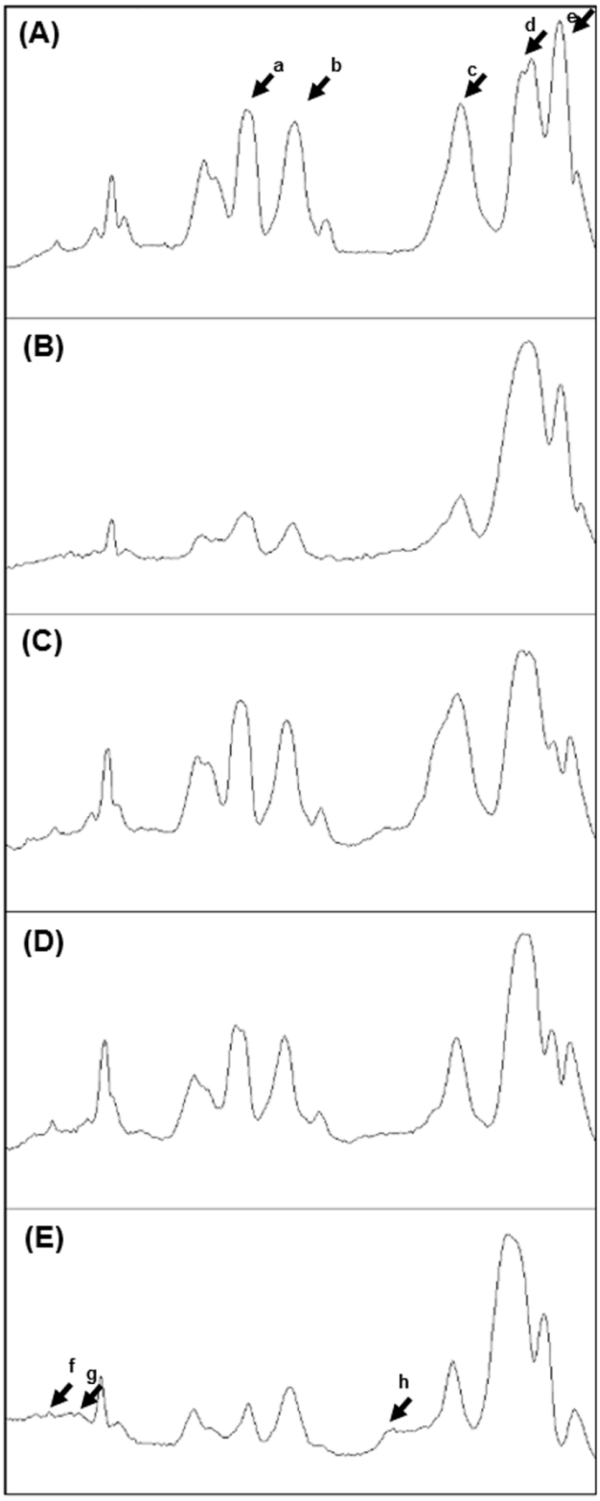


Figure S1. Densitometry analysis of Coomassie-stained SDS-PAGE protein profiles of the five WPC ingredient powders. (**A**) WPC Lac (batch 2), (**B**) WPC 35 (batch 3), (**C**) WPC HF1 (batch 2), (**D**) WPC HF2 (batch 1), and (**E**) WPC 80 (batch 2). a-h markers correspond to protein bands marked in Figure 1 (main text); d indicates β-Lg and e indicates α-Lac. Presented

densitometry analysis image is representative from one batch but values (Table S4) are calculated from the average of the three batches (or two batches for WPC 80).

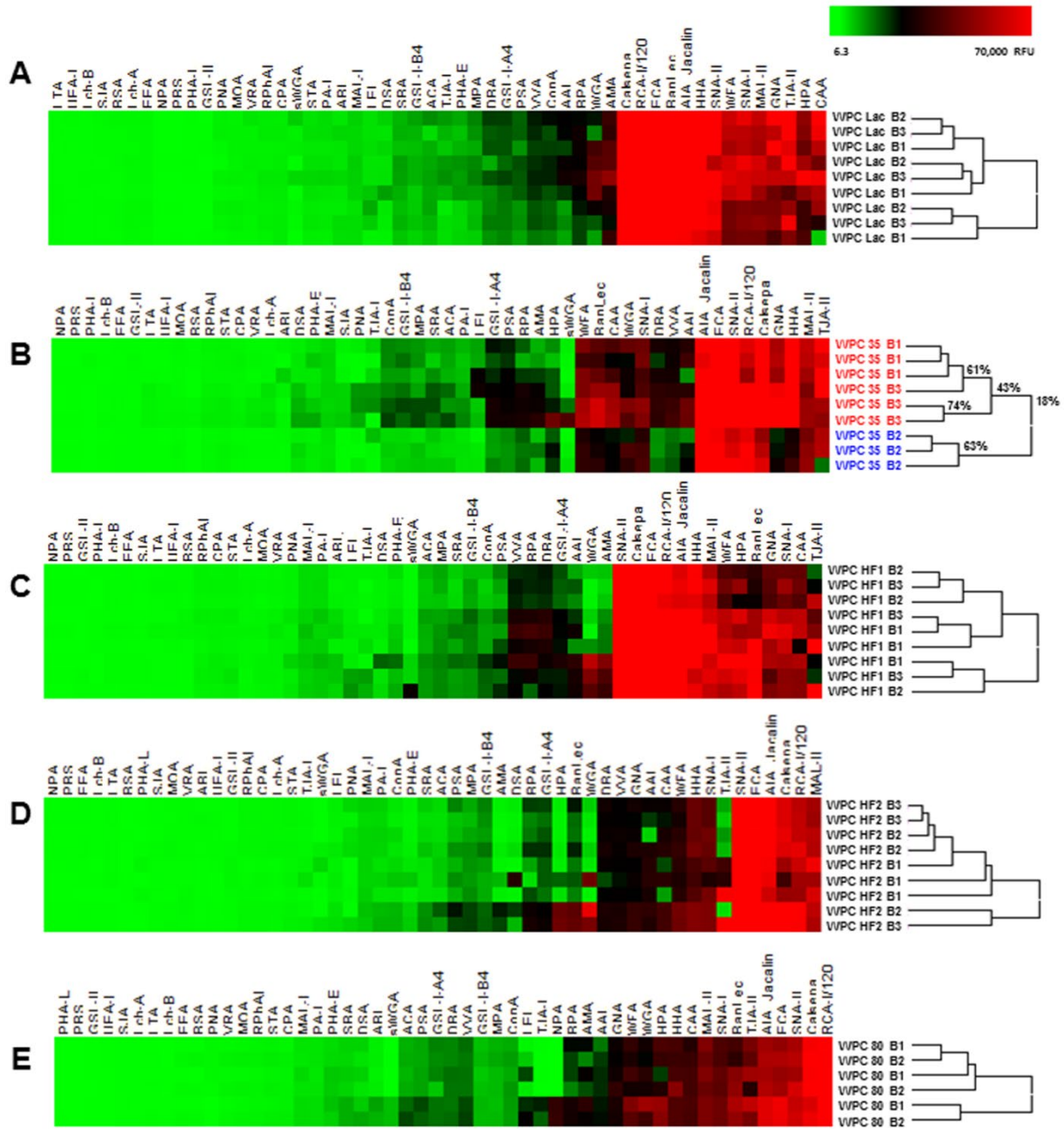


Fig. S2. Unsupervised clustering of lectin microarray profiles for the fluorescently-labelled production batches of the WPC ingredients. Each individual batch was incubated on lectin microarrays in triplicate. Normalised data were subjected to unsupervised, Euclidean distance, complete linkage clustering. (A) WPC Lac batches 1, 2 and 3, (B) WPC 35 batches 1, 2 and 3, where percentage similarity of pattern is indicated on the nodes, (C) WPC HF1 batches 1, 2 and 3, (D) WPC HF2 batches 1, 2 and 3, and (E) WPC 80 batches 1 and 2.

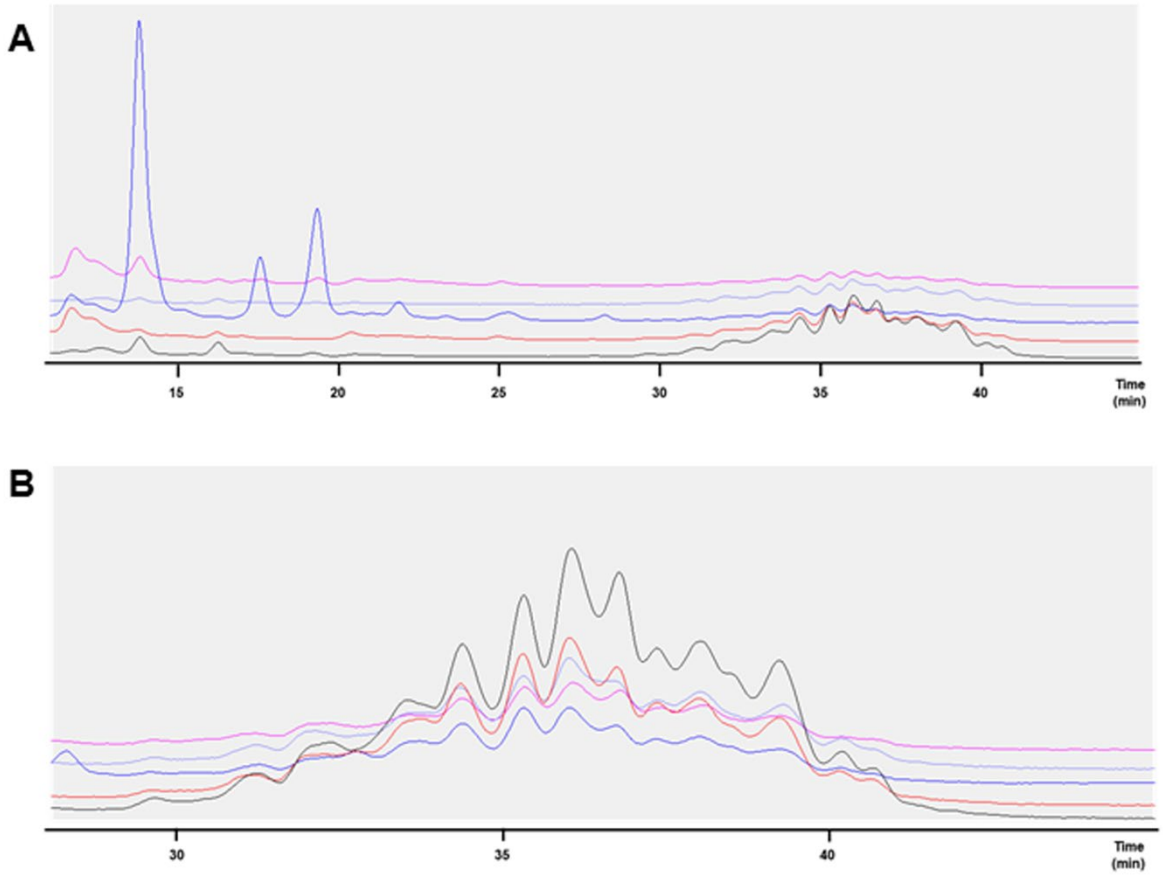


Fig. S3. Representative NP-HPLC chromatographs of 2AB-labelled N-linked glycans released from WPC 35 (blue), WPC 80 (pink), WPC Lac (black), WPC HF1 (red), WPC HF2 (purple) ingredient powders using the conventional PNGase F method. (A) Chromatographs from approximately 10-45 min. (B) Zoom-in on complex structures separation (approx. 25-48 min). Chromatographs are automatically rescaled by software upon zooming in.

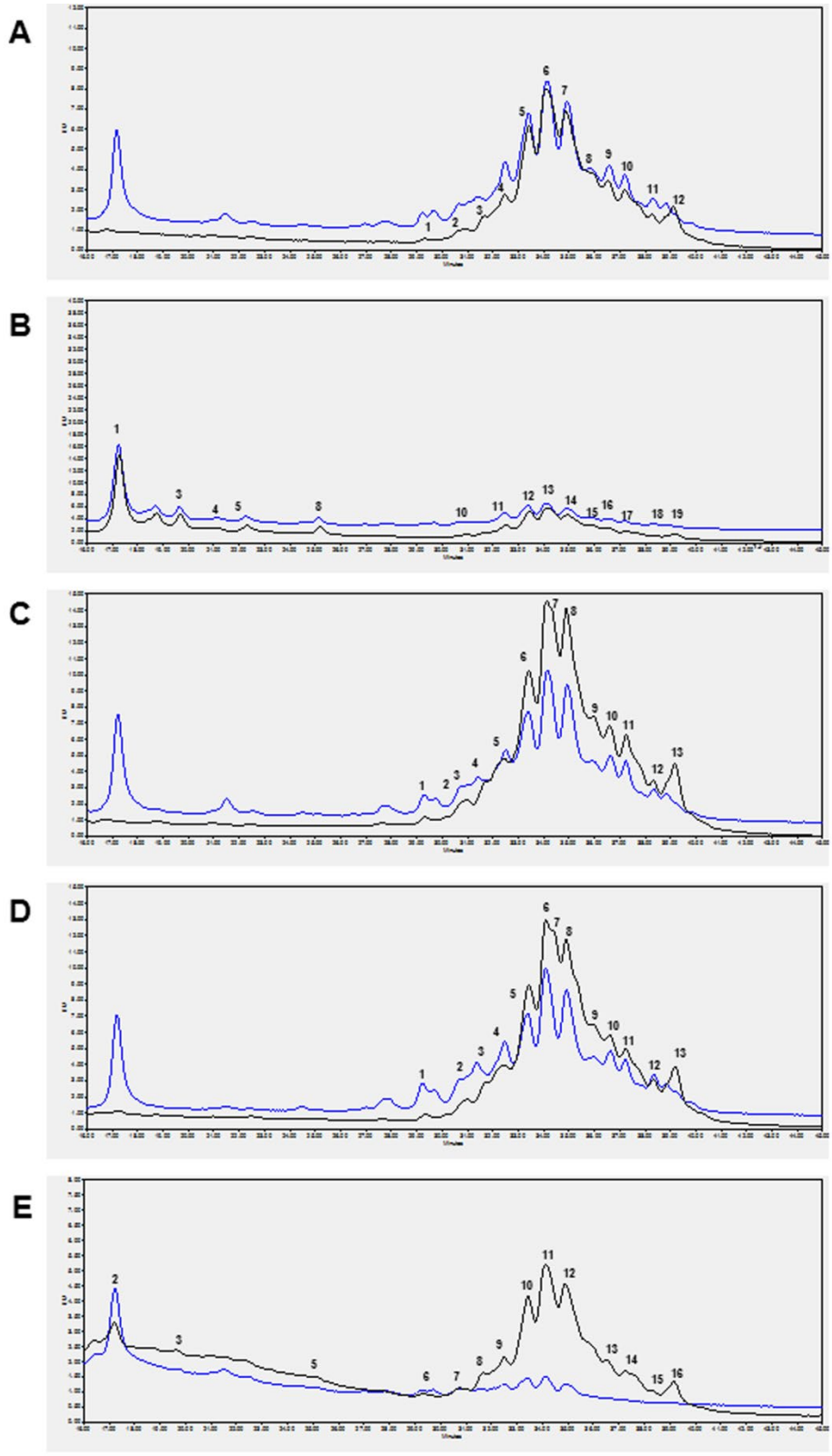


Fig. S4. Representative NP-HPLC chromatographs from 16-45 minutes of N-linked glycans released from WPC ingredient powders using the optimised method. (A) WPC Lac, (B) WPC 35, (C) WPC HF1, (D) WPC HF2, and (E) WPC 80. Black, non-digested; blue, mannosidase digested.

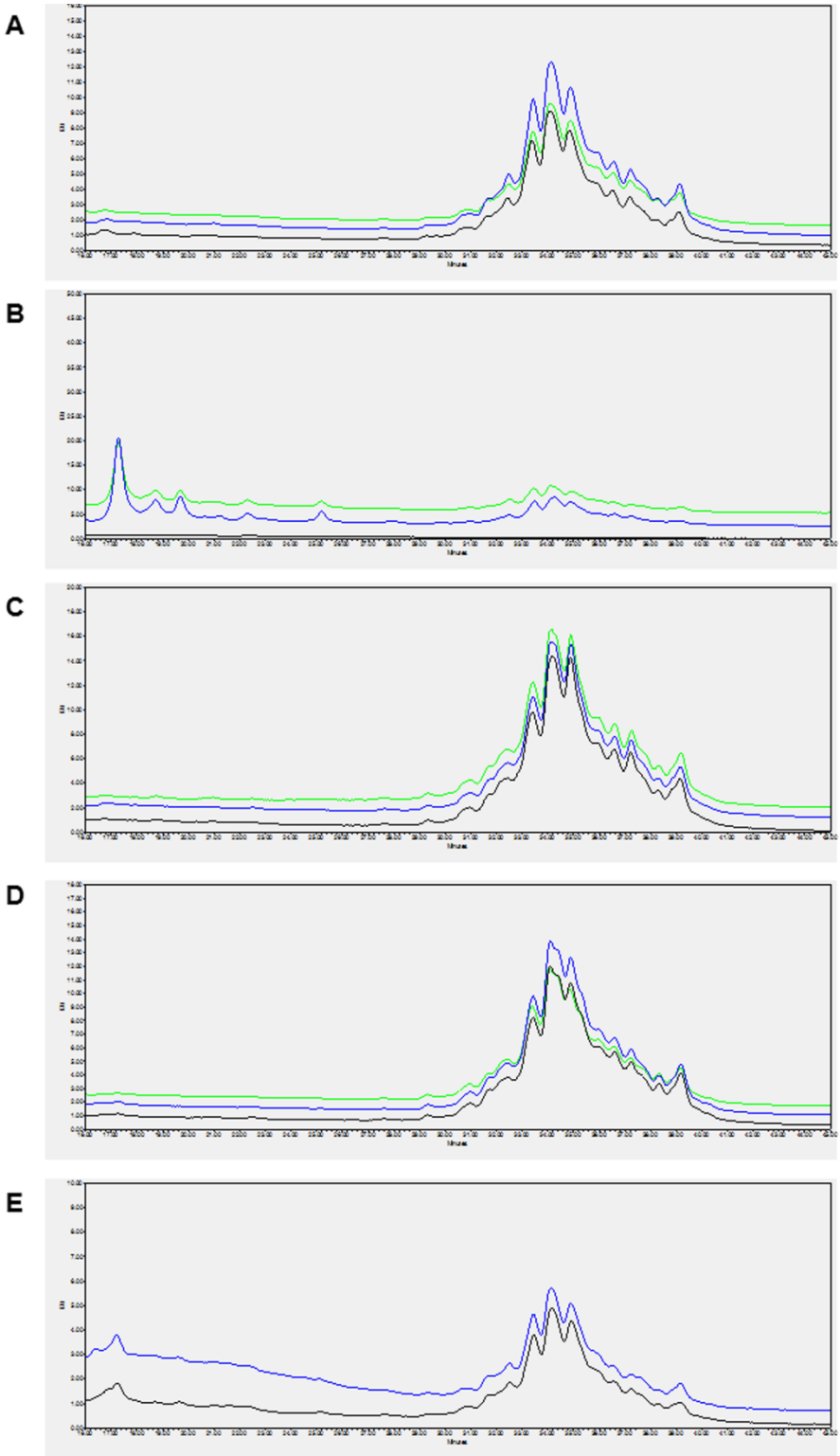


Fig. S5. Representative NP-HPLC chromatographs from 16-45 minutes of N-linked glycans from (A) WPC Lac batches 1 (black), 2 (blue) and 3 (green), (B) WPC 35 batches 2 (blue) and 3 (green), (C) WPC HF1 batches 1 (black), 2 (blue) and 3 (green), (D) WPC HF2 batch 1 (black), 2 (blue) and 3 (green), and (E) WPC 80 batches 1 (black) and 2 (blue).

Table S1. Sample code, description and average protein and fat percentages of the different WPC ingredient powders analysed provided by the manufacturers.

Sample code	Batch	Description	% protein (range)	% fat (range)
WPC 35	B1 B2 B3	WPC, reference	35 (34-40)	(2-2.5)
WPC 80	B1 B2	WPC, reference	80	6
WPC Lac	B1 B2 B3	α-Lac-enriched WPC The α-Lac used is precipitated from cheese or acid casein whey by calcium and contains elevated levels of phospholipids.	80	11
WPC HF1	B1 B2 B3	MFGM-enriched WPC	(71-74)	(17-20)
WPC HF2	B1 B2 B3	MFGM-enriched WPC with elevated levels of phospholipids	72	16

Table S2. Macronutrient composition of the five ingredient powders analysed in this work.

Sample	Protein as- is (g/100g)	CHO (g/100g)	Fat (g/100g)	Moisture (g/100g)	Ash (g/100g)
WPC Lac	78.2	2.70	10.5	5.69	2.86
WP HF1	72.0	4.20	17.8	4.20	1.78
WP HF2	70.1	8.00	15.0	5.00	1.93
WPC 80	74.9	9.90	6.00	6.82	2.44
WPC 35	35.6	50.9	2.80	3.49	6.31

Table S3. Lectins printed on the lectin microarray, their binding specificities, their simple print sugars (1 mM) and the supplying company. Binding specificity is reported recognition based on literature consensus or experimental evidence generated within our laboratory.

Abbreviation	Source	Species	Common name	General binding specificity	Print sugar	Supplier
AIA, Jacalin	Plant	*Artocarpus integrifolia*	Jack fruit lectin	Gal, Gal-β-(1,3)-GalNAc (sialylation independent)	Gal	EY Labs
RPbAI	Plant	*Robinia pseudoacacia*	Black locust lectin	Gal	Gal	EY Labs
SNA-II	Plant	*Sambucus nigra*	Sambucus lectin-II	Gal/GalNAc	Gal	EY Labs
SJA	Plant	*Sophora japonica*	Pagoda tree lectin	β-linked GalNAc	Gal	EY Labs
DBA	Plant	*Dolichos biflorus*	Horse gram lectin	GalNAc	Gal	EY Labs
SBA	Plant	*Glycine max*	Soy bean lectin	GalNAc	Gal	EY Labs
VVA	Plant	*Vicia villosa*	Hairy vetch lectin	GalNAc	Gal	EY Labs
BPA	Plant	*Bauhinia purpurea*	Camels foot tree lectin	GalNAc/Gal	Gal	EY Labs
WFA	Plant	*Wisteria floribunda*	Japanese wisteria lectin	GalNAc/sulfated GalNAc	Gal	EY Labs
HPA	Animal	*Helix pomatia*	Edible snail lectin	α-linked GalNAc	Gal	EY Labs
GSL-I-A4	Plant	*Griffonia simplicifolia*	Griffonia isolectin I A4	GalNAc	Gal	EY Labs
ACA	Plant	*Amaranthus caudatus*	Amaranthin	Sialylated/Gal-β-(1,3)-GalNAc	Lac	Vector Labs
ABL	Fungus	*Agaricus bisporus*	Edible mushroom lectin	Gal-β(1,3)-GalNAc, GlcNAc	Lac	EY Labs
PNA	Plant	*Arachis hypogaea*	Peanut lectin	Gal-β-(1,3)-GalNAc	Lac	EY Labs
GSL-II	Plant	*Griffonia simplicifolia*	Griffonia lectin-II	GlcNAc	GlcNAc	EY Labs
sWGA	Plant	*Triticum vulgaris*	Succinylated WGA	GlcNAc	GlcNAc	EY Labs
DSA	Plant	*Datura stramonium*	Jimson weed lectin	GlcNAc	GlcNAc	EY Labs
STA	Plant	*Solanum tuberosum*	Potato lectin	GlcNAc oligomers	GlcNAc	EY Labs
LEL	Plant	*Lycopersicum eculentum*	Tomato lectin	GlcNAc-β-(1,4)-GlcNAc	GlcNAc	EY Labs
BanLec	Plant	*Musa paradisciaca*	Banana lectin	Man/Glc in α-(1,3) linkage; laminaribose Glc-β-(1,3)-Glc	Man	Vector Labs
Calsepa	Plant	*Calystegia sepium*	Bindweed lectin	Man/Maltose	Man	EY Labs
NPA	Plant	*Narcissus pseudonarcissus*	Daffodil lectin	α-(1,6)-Man	Man	EY Labs
GNA	Plant	*Galanthus nivalis*	Snowdrop lectin	Man-α-(1,3)-	Man	EY Labs
HHA	Plant	*Hippeastrum hybrid*	Amaryllis agglutinin	Man-α-(1,3)-Man-α-(1,6)-	Man	EY Labs
Con A	Plant	*Canavalia ensiformis*	Jack bean lectin	Man, Glc, GlcNAc	Man	EY Labs
Lch-B	Plant	*Lens culinaris*	Lentil isolectin B	Man, core fucosylated, agalactosylated biantennary *N*-glycans	Man	EY Labs

Lch-A	Plant	*Lens culinaris*	Lentil isolectin A	Man/Glc	Man	EY Labs
PSA	Plant	*Pisum sativum*	Pea lectin	Man, core fucosylated trimannosyl *N*-glycans	Man	EY Labs
TJA-I	Plant	*Trichosanthes japonica*	Trichosanthes japonica agglutinin I	NeuAc-α-(2,6)-Gal-β-(1,4)-GIcNAc	Lac	Medicago
WGA	Plant	*Triticum vulgaris*	Wheat germ agglutinin	NeuAc/GlcNAc	GlcNAc	EY Labs
MAL-I	Plant	*Maackia amurensis*	Maackia agglutinin II	Sialic acid-α-(2,3)-Gal-β-(1,4)-GlcNAc; SO_4-3-Gal-β-(1,4)-GlcNAc	Lac	Vector Labs
MAL-II	Plant	*Maackia amurensis*	Maackia agglutinin II	Sialic acid-α-(2,3)-Gal-β-(1,3)-(±Sialic acid-α-(2,6)-)GalNAc; SO_4-3-Gal-β-	Lac	Vector Labs
SNA-I	Plant	*Sambucus nigra*	Sambucus lectin-I	Sialic acid-α-(2,6)-linked	Lac	EY Labs
PHA-L	Plant	*Phaseolus vulgaris*	Kidney bean leukoagglutinin	Tri- and tetraantennary β-Gal/Gal-β-(1,4)-GlcNAc	Lac	EY Labs
PHA-E	Plant	*Phaseolus vulgaris*	Kidney bean erythroagglutinin	Biantennary with bisecting GlcNAc,β-Gal/Gal-β-(1,4)-GlcNAc	Lac	EY Labs
RCA-I/120	Plant	*Ricinus communis*	Castor bean lectin I	Gal-β-(1,4)-GlcNAc	Gal	Vector Labs
AMA	Plant	*Arum maculatum*	Lords and ladies lectin	Gal-β-(1,4)-GlcNAc	Lac	EY Labs
CPA	Plant	*Cicer arietinum*	Chickpea lectin	Complex oligosaccharides	Lac	EY Labs
CAA	Plant	*Caragana arborescens*	Pea tree lectin	Gal-β-(1,4)-GlcNAc	Lac	EY Labs
ECA	Plant	*Erythrina cristagalli*	Cocks comb/coral tree lectin	Gal-β-(1,4)-GlcNAc oligomers	Lac	EY Labs
TJA-II	Plant	*Trichosanthes japonica*	Trichosanthes japonica agglutinin II	Fuc-α-(1,2)-Gal-β-(1,4)-GlcNAc, terminal β-linked GalNAc	Lac	Medicago
AAL	Fungi	*Aleuria aurantia*	Orange peel fungus lectin	Fuc-α-(1,6)-linked, Fuc-α-(1,3)-linked	Fuc	Vector Labs
LTA	Plant	*Lotus tetragonolobus*	Lotus lectin	Fuc-α-(1,3)-linked	Fuc	EY Labs
UEA-I	Plant	*Ulex europaeus*	Gorse lectin-I	Fuc-α-(1,2)-Gal	Fuc	EY Labs
PA-I	Bacteria	*Pseudomonas aeruginosa*	Pseudomonas lectin	Terminal α-linked Gal, Gal derivatives	Gal	EY Labs
EEA	Plant	*Euonymous europaeus*	Spindle tree lectin	Terminal α-linked Gal	Gal	EY Labs
GSL-I-B4	Plant	*Griffonia simplicifolia*	Griffonia/Bandeiraea lectin I B4	Terminal α-linked Gal	Gal	EY Labs
MPA	Plant	*Maclura pomifera*	Osage orange lectin	Terminal α-linked Gal	Gal	EY Labs
VRA	Plant	*Vigna radiata*	Mung bean lectin	Terminal α-linked Gal	Gal	EY Labs
MOA	Fungus	*Marasmius oreades*	Fairy ring mushroom lectin	Terminal α-linked Gal	Gal	EY Labs
PBS	N/A	N/A	Phosphate buffered saline	N/A	N/A	Sigma
BSA	Animal	N/A	Bovine serum albumin	N/A	N/A	Sigma

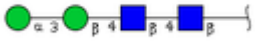
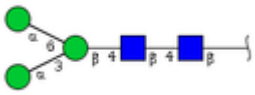
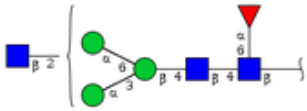
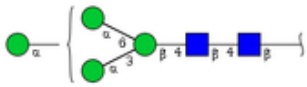
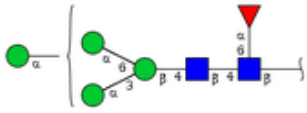
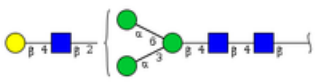
Table S4. Integrated densitometry analysis of Coomassie-stained SDS-PAGE protein profiles of the five WPC ingredients using ImageJ. Relative percentage is an average from the three batches (or two batches for WPC 80) for each ingredient.

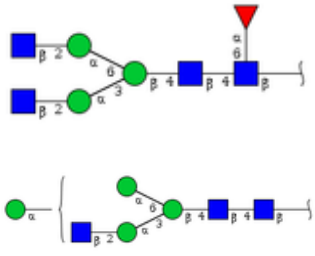
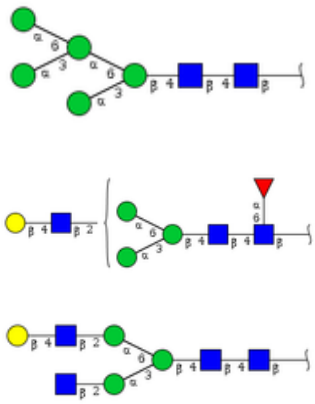
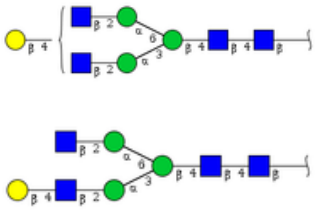
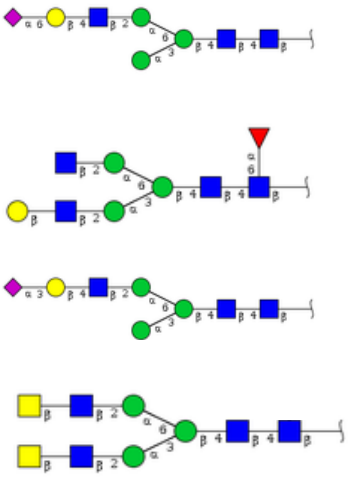
Peaks of interest	WPC Lac	WPC 35	WPC HF1	WPC HF2	WPC 80
f	0.48	0.29	0.42	0.56	
g	0.61	0.27	0.46	0.47	
	2.48	2.66	3.45	4.82	3.19
	0.74	0.77	0.25	0.33	0.83
	8.98	2.00	9.68	8.55	6.09
a	12.97	8.57	13.38	14.20	4.93
b	15.64	6.41	13.35	13.30	12.08
	1.18	0.37	1.44	1.48	
h					2.20
c	23.57	11.03	28.01	16.38	10.82
d	17.83	52.73	23.88	32.06	44.87
e	14.76	13.71	0.52	2.11	8.76
	0.77	1.20	5.17	5.74	6.24

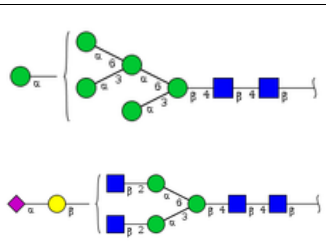
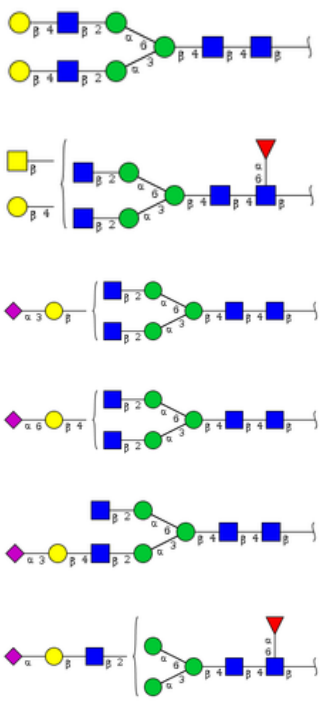
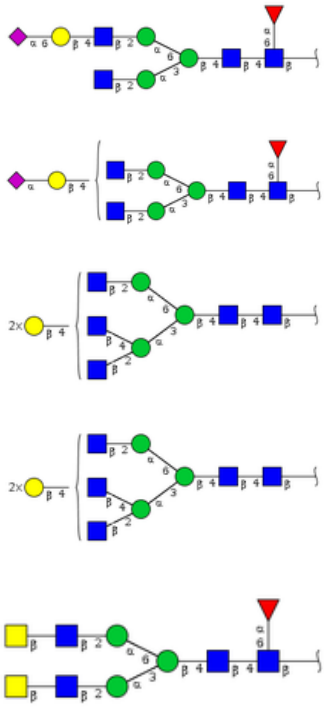
Table S5. Peak GU values, relative quantification pre- (% area) and post-mannosidase digestion (%D.A.), and assigned potential N-linked glycan structures from the five WPC ingredients.

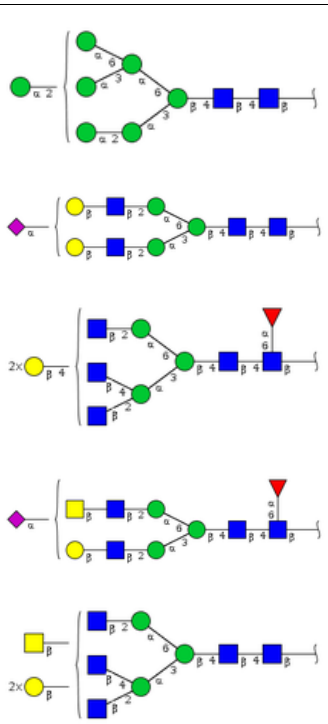
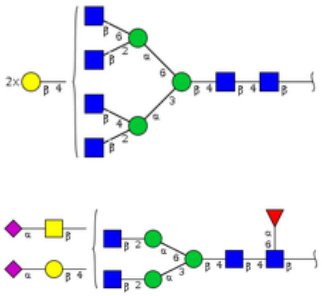
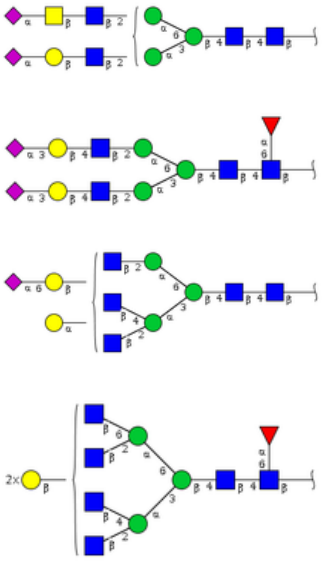
WPC 35	GU	% Area	% D.A.	WPC 80	GU	% Area	% D.A.	WPC lac	GU	% Area	% D.A	WPC HF1	GU	% Area	% D.A.	WPC HF2	GU	% Area	% D.A	Potential Structures
				1	2.52	0.69														
1	2.68	22.79	26.22	**2**	2.67	2.78	45.27		2.66		8.12		2.66		8.75		2.66		7.69	M1
2	2.95	7.73																		
3	3.13	5.40		**3**	3.11	0.24														
4	3.32	0.97																		
5	3.42	1.35	2.48																	M2
6	3.66	2.82		**4**	3.66	1.38														
7	4.18	0.10																		
8	4.31	1.55		**5**	4.29	0.35														F(3)M2/M3
9	5.01	0.47																		
				6	5.44	0.15		**1**	5.43	0.18	1.12	**1**	5.42	0.56	1.75	**1**	5.42	0.58	2.63	A2/F(6)A1/M4
												2	5.70	0.31						F(6)M4/A1G(4)1
10	5.96	1.63	1.39	**7**	5.95	1.38		**2**	5.95	1.52		**3**	5.95	2.32		**2**	5.94	2.57		F(6)A2/M4A1
				8	6.21	2.26		**3**	6.20	1.83		**4**	6.21	2.21		**3**	6.19	2.12		M5/F(6)A1G(4)1/A2[6]G(4)1
11	6.49	6.56		**9**	6.49	6.87	0.58	**4**	6.48	6.40		**5**	6.46	6.10		**4**	6.45	6.98		A2G(4)1/A2[6]G(4)1
12	6.85	9.56	8.75	**10**	6.84	14.94	12.06	**5**	6.83	14.78	13.2	**6**	6.82	14.12	13.01	**5**	6.81	14.32	11.3	A2GalNAc2/F(6)A2[3]G(4)1/A1[6]G(4)1S(3)1/A1[6]G(4)1S(6)1
13	7.11	12.68	9.97	**11**	7.11	21.34	9.95	**6**	7.10	22.01	16.6	**7**	7.10	20.15	15.49	**6**	7.07	13.60		M6/M6D1/M6D2/A2G1S1
																7	7.17	8.77		A2G(4)2/A2G1S(3)1/F(6)A2G(4)1GalNAc1/A2G4)1S(6)1/A2[3]G(4)1S(3)1/F(6)A1G1S1
14	7.44	10.39	8.86	**12**	7.42	28.84	9.49	**7**	7.41	22.45	15.7	**8**	7.41	22.30	15.93	**8**	7.39	20.80	14.4	F(6)A2G(4)GlcNAc1/F(6)A2[6]G(4)1S(6)1/F(6)A2G(4)1S1/A3G(4)2/F(6)A3G1
15	7.77	4.66	4.19					**8**	7.85	5.46		**9**	7.87	6.23	5.62	**9**	7.88	6.60	6.24	M7/A2G2S1/F(6)A3G(4)2/F(6)A2G1GalNAc1S1/A3G2GalNAc1
16	8.16	3.40		**13**	8.16	5.14		**9**	8.16	6.74	6.26	**10**	8.17	6.56	5.9	**10**	8.16	6.43	5.65	A4G(4)2/F(6)A2G(4)1S1GalNAc1S1
17	8.50	4.08		**14**	8.49	7.33		**10**	8.47	9.22	6.63	**11**	8.48	9.66	6.57	**11**	8.46	7.87	4.93	A2G1GalNAc1S2/F(6)A2G(4)2S(3,3)2/A3G2S(6)1/F(6)A4G2
18	9.04	1.06		**15**	9.04	1.38		**11**	9.01	2.31		**12**	9.03	2.49	2.39	**12**	9.03	2.75		F(6)A2G(4)2S(6,6)2/F(6)A2G(4)2S2/F(6)A2G2Ga1S1/A3G3S1
19	9.50	2.79		**16**	9.50	4.94		**12**	9.47	6.98	3.05	**13**	9.49	7.00	2.87	**13**	9.48	6.60	4.46	M9/F(6)A2G2AS(6)1

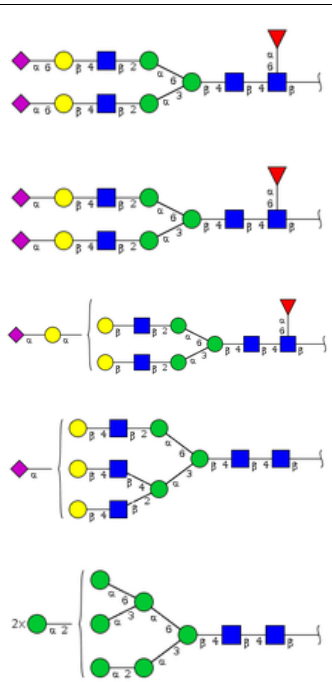
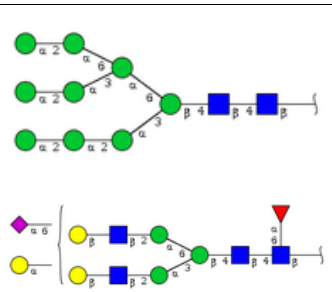
Table S6. Predicted structures present in WPC 35, WPC 80, WPC Lac, WPC HF1 and WPC HF2. n.d. not determined. Structure notation is according to the Oxford notation.

	RT (min)	GU	Structure	Structure notation
1	16.4	2.52	n.d.	
2	17.23	2.66		M1
3	18.73	2.95	n.d.	
4	19.65	3.11	n.d.	
5	20.67	3.32	n.d.	
6	21.16	3.42		M2
7	22.32	3.66		LNnT
8	24.63	4.18	n.d.	
9	25.14	4.30	 	DSL M3
10	27.86	5.01	n.d.	
11	29.31	5.42	  	A2 F(6)A1 M4
12	30.19	5.71	 	F(6)M4 A1G(4)1

13	30.95	5.95		F(6)A2 M4A1
14	31.70	6.20		M5 F(6)A1G(4) A2[6]G(4)1
15	32.47	6.47		A2G(4)1 A2[3]G(4)1
16	33.42	6.83		A1[6]G(4)1S(6)1 F(6)A2[3]G(4)1 A1[6]G(4)1S(3)1 A2GalNAc2

17	34.11	7.09		M6 A2G1S1
18	34.32	7.17		A2G(4)2 F(6)A2G(4)1GalNAc1 A2G1S(3)1 A2G(4)1S(6)1 A2[3]G(4)1S(3)1 F(6)A1G1S1
19	34.88	7.41		F(6)A2[6]G(4)1S(6)1 F(6)A2G(4)1S1 A3G(4)2 F(6)A3G1 F(6)A2GalNAc2

20	35.87	7.84		M7 A2G2S1 F(6)A3G(4)2 F(6)A2G1GalNAc1S1 A3G2GalNAc1
21	36.55	8.16		A4G(4)2 F(6)A2G(4)1S1GalNAc1S1
22	37.23	8.48		A2G1GalNAc1S2 F(6)A2G(4)2S(3,3)2 A3G2S(6)1 F(6)A4G2

23	38.31	9.03		F(6)A2G(4)2S(6,6)2 F(6)A2G(4)2S2 F(6)A2G2Ga1S1 A3G3S1 M8
24	39.16	9.48		M9 F(6)A2G2Ga1S(6)1