International Dairy Journal 129 (2022) 105340

Contents lists available at ScienceDirect

International Dairy Journal

journal homepage: www.elsevier.com/locate/idairyj

Identifying glycation hot-spots in bovine milk proteins during production and storage of skim milk powder

Inge Gazi [a, b, *], Vojtech Franc [a, b], Sem Tamara [a, b], Martine P. van Gool [c], Thom Huppertz [c, d], Albert J.R. Heck [a, b]

[a] Biomolecular Mass Spectrometry and Proteomics, Bijvoet Centre for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Padualaan 8, 3584, CH, Utrecht, the Netherlands
[b] Netherlands Proteomics Centre, Padualaan 8, 3584, CH, Utrecht, the Netherlands
[c] FrieslandCampina, Stationsplein 4, 3818, LE, Amersfoort, the Netherlands
[d] Department of Agrotechnology and Food Sciences, Wageningen University, Bornse Weilanden 9, 6708, WG, Wageningen, the Netherlands

ARTICLE INFO

Article history:
Received 29 September 2021
Received in revised form
12 January 2022
Accepted 24 January 2022
Available online 2 February 2022

ABSTRACT

We investigated protein glycation in a complex milk system under controlled conditions representative of real-life consumer products, analysing intermediate and final products from skim milk powder production, and aged powder samples. We combined protein-centric LC-MS(/MS) with peptide-centric multi-protease LC-MS/MS focusing on the six most abundant bovine milk proteins. This strategy resulted in the identification of glycated proteoforms and of the extent of glycation per protein, high protein sequence coverage, and identification and relative occupancy of the glycation sites. We identified new glycation hot-spots additionally to the ones already described in literature. Primary sequence motif analysis revealed that glycation hot-spots were preceded N-terminally by a stretch rich in basic amino acids, and followed C-terminally by a stretch enriched in aliphatic and hydrophobic amino acids. Our study considerably extends the current understanding of milk protein glycation, discussing glycation hot-spots and their localisation in relation to the primary sequences and higher-order protein structures.

1. Introduction

Heat treatment of milk is imperative for rendering it safe for consumption, specifically from the perspective of microbiological safety (Boor, Wiedmann, Murphy, & Alcaine, 2017). However, an unintended consequence of heat treatment is the series of chemical reactions, collectively known as the Maillard reaction (MR). This process starts with protein glycation and eventually leads to advanced glycation end products (van Boekel, 1998). Glycated proteins exhibit altered digestibility and reduced bioavailability of the essential amino acid lysine (van Lieshout, Lambers, Bragt, & Hettinga, 2020), whereas advanced MR products have been linked with increased immunogenicity, allergenicity, inflammation, and oxidative stress (Baye et al., 2017; Sharma & Barone, 2019; Teodorowicz, Van Neerven, & Savelkoul, 2017; Toda, Hellwig, Henle, & Vieths, 2019). Detailed investigation of glycation in milk systems is thus important to understand the reaction mechanism and to prevent or mitigate it.

Glycation of milk proteins and the advancement of the MR is most frequently analysed using MR indicators such as available lysine or furosine in the early stages of the reaction, 5-hydroxymethyl-2-furfuraldehyde (HMF) or 2-furfuraldehyde (furfural) in the intermediate stages, and browning index in the advanced stages of the reaction (Aalaei, Rayner, & Sjöholm, 2019; Gómez-Narváez, Perez-Martinez, & Contreras-Calderon, 2019). Alternatively, mass spectrometry (MS) of the six most abundant intact milk proteins, i.e., α_{S1}-, α_{S2}-, β- and κ-casein, α-lactalbumin and β-lactoglobulin (Arena, Renzone, D'Ambrosio, Salzano, & Scaloni, 2017; Oliver, 2011; Siciliano, Mazzeo, Arena, Renzone, & Scaloni, 2013), can be employed to analyse glycated proteins. This latter approach monitors mass shifts induced by adding glycans on proteins, revealing the overall extent of glycation, without identifying glycation sites and their occupancy.

Glycation site localisation in the protein structure is important for identifying newly-formed immunogenic and allergenic epitopes. A peptide-centric mass spectrometry approach allows site-specific

* Corresponding author.
E-mail address: i.gazi@uu.nl (I. Gazi).

https://doi.org/10.1016/j.idairyj.2022.105340

characterisation and complements MS on intact proteins (Franc, Zhu, & Heck, 2018; Tamara, Franc, & Heck, 2020; van de Waterbeemd et al., 2018; Yang et al., 2016). Trypsin is the gold standard protease in peptide-centric MS (Tsiatsiani & Heck, 2015), cleaving exclusively C-terminal of arginine and lysine residues (Olsen, Ong, & Mann, 2004). Glycation of these two residues hampers cleavage by trypsin, resulting in suboptimal digestion and hindering downstream LC-MS/MS analysis (Deng, Wierenga, Schols, Sforza, & Gruppen, 2017). Nevertheless, the greater part of published peptide-centric MS research on milk protein glycation still relies on tryptic digestion, despite this decreased performance of the enzyme towards glycated proteins. To gain a holistic picture of the extent of glycation, localisation and site occupancy during the processing and storage of skim milk powder, we combined here protein- and multi-protease peptide-centric MS approaches. This strategy resulted in high protein sequence coverage irrespective of the extent of glycation, overcoming the limitations of using solely trypsin.

Apart from the choice of protease in the peptide-centric approach, the selected dairy system to investigate and the experimental design may also lead to biased results. Studies focusing on heat- and storage-induced changes occurring in model systems of isolated milk proteins in combination with reducing sugars (Cardoso, Wierenga, Gruppen, & Schols, 2018; Carulli, Calvano, Palmisano, & Pischetsrieder, 2011; Cattaneo, Stuknytė, Masotti, & De Noni, 2017; Chevalier, Chobert, Dalgalarrondo, Choiset, & Haertlé, 2002; Gómez-Narváez et al., 2019; Leiva, Naranjo, & Malec, 2017) only consider a few isolated factors, e.g., protein and carbohydrate concentration, and type of carbohydrate. Due to differences in sample complexity and processing conditions, findings from model systems do not necessarily translate well to consumer dairy systems. Glycation studies conducted on off-the-shelf commercial products (Birlouez-Aragon et al., 2004; Delatour et al., 2009; Hegele, Buetler, & Delatour, 2008; Milkovska-Stamenova & Hoffmann, 2016a, 2016b; Renzone, Arena, & Scaloni, 2015) generate results directly representative of the consumer dairy products. However, the difficulty in interpreting results from these studies comes from the fact that the starting ingredients and processing and storage conditions are all unknown, and may differ between the same types of products from different brands.

Our study was designed for the in-depth investigation of protein glycation during the production and storage of skim milk powder. The production process was carried out on pilot scale representative of industrial processing to overcome the previously described limitations. All processing and storage steps were performed under controlled conditions. All intermediate, final and aged samples belonged to the same batch. The extent of glycation was monitored in a site-specific manner for the six most abundant milk proteins. We observed the gradual increase of glycation and substantial residue-specific differences, and identified glycation hot-spots in these six proteins. We revealed a generic preferential sequence motif of glycated lysine residues, thus expanding the current understanding of milk protein glycation with new insights.

2. Materials and methods

2.1. Reagents

All chemical and biochemical reagents were purchased from Sigma–Aldrich (Germany), unless otherwise specified.

2.2. Milk samples

Skim milk powder was produced on pilot scale at the FrieslandCampina Innovation Centre in Wageningen, The Netherlands, following the flowchart shown in Fig. 1A. The batch of samples was prepared from mature bulk milk from Dutch Friesian Holstein cows. Since these samples are representative of the average Dutch bovine milk, no additional biological replicates were analysed. Briefly, skim milk was thermised and standardised to 35% (m m^{-1}) protein in dry matter. The standardised milk was preheated for 60 s at 108 °C. The preheated milk was further concentrated to 48–50% (m m^{-1}) dry matter, followed by spray-drying (inlet air temperature = 205 °C; outlet air temperature = 90 °C) resulting in <0.2 water activity in the powder. The skim milk powder was sealed in tin cans that were further subjected to a storage trial at either 20 or 40 °C for 43, 77, and 159 d. Aliquots of fresh and aged powders were stored at −20 °C until further analysis.

2.3. Gel electrophoresis and analysis of blocked lysine residues

All intermediate, final and aged products shown in Fig. 1A were subjected to sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) under both reducing (with dithiothreitol, DTT) and non-reducing (without DTT) conditions. The samples were treated with XT sample buffer, loaded onto 26-well 12% CriterionTM XT Bis-Tris Precast Gels, and electrophoresis was run in XT MOPS running buffer, all purchased from Bio-Rad (Veenendaal, The Netherlands). Precision Plus Protein™ Dual Color Standards (Bio-Rad, The Netherlands) were run on the gels in parallel with the samples for protein size reference. The gels were stained with GelCode Blue (Thermo Fisher Scientific, Landsmeer, The Netherlands). The detailed experimental procedure is described in the Supplementary material.

Blocked (i.e., glycated) lysine was determined for the thermised and preheated milk and for the fresh and aged (159 d at 40 °C) skim milk powders following previously described procedures by Nyakayiru et al. (2020).

Acid hydrolysis of Amadori products leads to the formation of furosine, which is not endogenous to the dairy products. Lysine and furosine were quantified using ion-pair RP-HPLC following acid hydrolysis (Delgado, Corzo, Santa-María, Jimeno, & Olano, 1992). Krause, Knoll, and Henle (2003) determined a molar yield of 34% lactulosyllysine to furosine, which was used for the conversion of the measured furosine contents. Blocked lysine was calculated as lactulosyllysine expressed as a percentage of the total lysine contents of the sample.

2.4. Protein-centric mass spectrometry

The thermised and preheated milk and the fresh and aged (159 d at 40 °C) milk powders were selected for analysis by mass spectrometry. The powders were reconstituted in MilliQ at 10% (m m^{-1}) powder in solution. The samples were denatured and reduced according to a method modified from Bondt et al. (2021). LC-MS and MS/MS analyses were performed using a Thermo Scientific Vanquish Flex UHPLC system (Thermo Fisher Scientific, Germering, Germany) connected online to an Orbitrap Fusion Lumos Tribrid mass spectrometer (Thermo Fisher Scientific, San Jose, CA, USA) operated in intact protein mode and at low pressure. Reversed-phase separation was achieved with a MAbPac RP 2.1 mm × 50 mm column (Thermo Fisher Scientific, Germany). Each sample was analysed with a triplet of complementary methods comprising full MS, MS/MS with electron transfer dissociation (ETD) and MS/MS with electron transfer dissociation supplemented by higher-energy collisional dissociation (EThcD). No additional technical replicates were performed. LC-MS raw data were deconvoluted with the BioPharma Finder 3.2 Software using the Xtract algorithm (Thermo Fisher Scientific, USA). The database for searching the LC-MS/MS results was generated based on the six most abundant bovine milk proteins downloaded in XML format

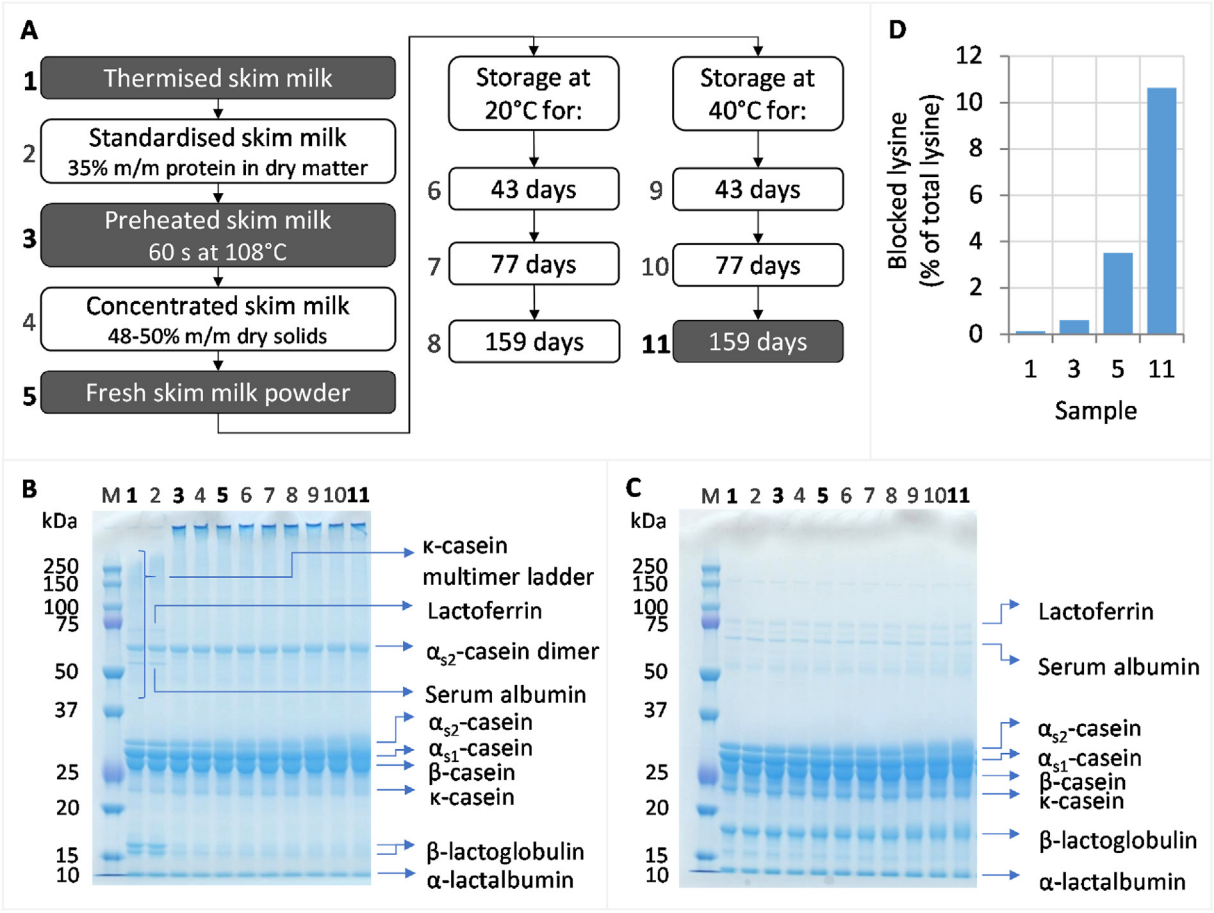

Fig. 1. Physical-chemical changes in the major milk proteins during production and storage of skim milk powder analysed by SDS-PAGE. (A) Flowchart depicting the eleven samples analysed, namely the intermediate and final products of the skim milk powder production process, and the aged powders obtained by storage at 20 and 40 °C, respectively. (B) and (C) SDS-polyacrylamide gel electrophoretograms of all samples shown in the flowchart (A) analysed under non-reducing (B) and under reducing (C) conditions; M = molecular mass marker. (D) Total percentage of blocked lysine residues determined as lactosylated lysine expressed as a percentage of total lysine in samples 1, 3, 5 and 11 as annotated in flowchart (A).

from UniProtKB (accession numbers P00711, P02754, P02662, P02663, P02666 and P02668) and imported from the XML format file into ProSightPC. The LC-MS/MS results were automatically searched in Thermo Proteome Discoverer (v 2.4.0.305) using the ProSightPD nodes for High–High experimental workflows. A library of proteoforms and corresponding monoisotopic masses was generated to include the most commonly occurring genetic variants and expected phosphorylation states of the major bovine milk proteins in Dutch Holstein-Friesian cows. For each basic proteoform, the masses of glycated proteoforms were generated in a range of 1–12 hexose residues at increments of 1 hexose residue. Proteoforms were annotated based on LC-MS/MS database search identification and matching between the experimental and theoretical masses. Sample preparation, LC-MS and MS/MS analyses, database generation, database search, library generation and data analysis are described in detail in the Supplementary material.

2.5. Peptide-centric mass spectrometry

The four samples mentioned above were also analysed by a peptide-centric mass spectrometry approach. The samples were denatured, reduced and alkylated, followed by proteolytic digestion with trypsin (Promega, Madison, WI, USA), chymotrypsin, GluC, AspN (latter three purchased from Roche, Mannheim, Germany), and combinations thereof as described in Table 1. Since each sample was therefore analysed from a total of seven different and complementary proteolytic digests, no additional technical replicates were performed. Following proteolytic digestion, the resulting peptides were extracted by solid-phase extraction using Oasis HLB 96-well plates (Waters, Etten-Leur, The Netherlands). The samples were analysed using an Agilent 1290 Infinity HPLC system (Agilent Technologies, Middelburg, The Netherlands) coupled online to a Q Exactive HF hybrid quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) according to a method modified from Zhu et al. (2020). The samples were separated by reversed-phase chromatography using in-house packed trapping (2 cm length, 100 μm inner diameter; ReproSil-

Table 1
Protease combinations and digestion conditions used for hydrolysing the milk proteins.[a].

Digestion	Step 1		Step 2	
	Enzyme	Incubation	Enzyme	Incubation
1	Trypsin	37 °C, overnight	–	–
2	Chymotrypsin	RT, overnight	–	–
3	GluC	RT, overnight	–	–
4	AspN	37 °C, overnight	–	–
5	Chymotrypsin	RT, 4 h	Trypsin	37 °C, overnight
6	GluC	RT, 4 h	Trypsin	37 °C, overnight
7	AspN	37 °C, 4 h	Trypsin	37 °C, overnight

[a] Abbreviation: RT, room temperature. Each enzyme was added at a concentration of 1 μg enzyme μL^{-1} denatured milk sample.

Pur C18-AQ, 3 μm; Dr. Maisch GmbH, Ammerbuch, Germany) and analytical columns (50 cm length, 50 μm inner diameter; Poroshell 120 EC-C18, 2.7 μm; Agilent Technologies, The Netherlands).

LC-MS/MS raw data were searched with Thermo Proteome Discoverer V2.4 (Thermo Fisher Scientific, Germany) using the Mascot V2.6 (Matrix Science, Boston, MA, USA) search engine. To ensure high identification rate, the modifications taken into consideration in the database search included glycation with mono- and dihexose of protein N-termini, lysine and arginine residues, phosphorylation of serine and threonine residues, oxidation of methionine residues, cyclisation of N-terminal glutamine residues to pyroglumatic acid, and fixed carbamidomethylation of cysteine residues. The data were carefully post-processed, including filtering of unreliable identifications and normalisation across digests and injections, prior to combining into a comprehensive picture. Glycation site occupancy was calculated as the sum of abundances of all peptide isoforms covering a glycated site expressed as a percentage of the total sum of abundances of all peptide isoforms covering the respective site, whether glycated or not. Sample preparation, LC-MS/MS analysis, database search and data processing are described in detail in the Supplementary material.

2.6. Glycated lysine motif analysis

Glycation motif analysis was performed for the six most abundant milk proteins on the reference sequences from UniProtKB. The signal peptides were excluded to focus the analysis on the mature proteins present in the milk. The glycated sites chosen for the analysis were lysine residues, whereby the glycation site occupancy was detected at >5% of relative abundance in the aged milk powder. The sequence logos were constructed for the glycated and non-glycated lysine residues in the six most abundant milk proteins, using standard position weighted matrices (Hertz & Stormo, 1999). For both the glycated and non-glycated lysine residues, their sequence motifs were extracted for the ten amino acids preceding and following them. The sequence motif of the non-glycated lysine residues was subsequently subtracted from that of the glycated ones to highlight the differences better.

3. Results

Our study aimed to investigate protein glycation in consumer products, analysing samples taken during the processing of skim milk into skim milk powder on pilot scale, and during subsequent aging of the powder. To study protein glycation under these conditions, we analysed a batch of 11 different samples in parallel, as schematically summarised in Fig. 1A.

3.1. Gel electrophoresis and determination of the percentage of blocked lysine residues

Samples were analysed by SDS-PAGE to obtain a comprehensive overview of the steps where the proteins undergo changes, and of the types of changes that occur. Images of the SDS-polyacrylamide gels of the samples summarised in Fig. 1A are shown in Fig. 1B and C. High molecular mass disulphide-linked protein aggregates were formed due to preheating the milk at 105 °C for 60 s; these aggregates were too large to enter the gel and therefore were visible as intense bands at the top of the non-reducing gel (Fig. 1B). The aggregates were not observed in the reducing gel (Fig. 1C) where the individual protein bands became visible, further confirming that the nature of the aggregates was indeed disulphide-linked and not that of other covalent linkages.

Another noticeable change observed in the gels is the broadening of the protein bands (Fig. 1B and C). This started to occur in the preheated milk and increased with subsequent processing, and particularly during storage of the powder. The broadening of the protein bands likely indicates that the original proteins have differentiated into a series of more diverse proteoforms with altered migration times due to process- and storage-induced chemical modifications.

Based on the SDS-PAGE results (Fig. 1B and C), four key samples were chosen for further analyses, i.e., the thermised milk (sample 1), the preheated milk (sample 3), the fresh skim milk powder (sample 5), and the aged skim milk powder (sample 11; 159 d at 40 °C).

The broadening of the protein bands on SDS-PAGE was presumed to be primarily caused by glycation of the proteins with lactose, i.e., lactosylation. The four samples specified above were analysed for the percentage of blocked lysine residues to confirm whether protein glycation was indeed responsible for the observed changes. The results from the blocked lysine analyses are shown in Fig. 1D, confirming that lactosylation occurred during processing and storage. Virtually no blocked lysine was detected in the thermised milk. A low level of blocked lysine (<1% of total lysine) was detected in the preheated milk following heat treatment at 108 °C for 60 s. Further heat load during concentration of the milk and spray-drying led to the increase of blocked lysine to 3.5% in the fresh skim milk powder. Our experimental storage conditions (water activity <0.2 and temperature of 20° or 40 °C) ensured that the lactose in the milk powders remained below its glass transition temperature (T_g; Schuck et al., 2005). Despite the limited molecular mobility below T_g and thus slower glycation reaction kinetics (Gonzales, Naranjo, Leiva, & Malec, 2010), the level of blocked lysine further increased to 10.6% of total lysine during storage of the powder at 40 °C for 159 d.

3.2. Protein-centric mass spectrometry

We further analysed the thermised and preheated milk, and the fresh and aged (159 d at 40 °C) skim milk powders by protein-centric mass spectrometry. Each sample was analysed by a triplet of one LC-MS and two LC-MS/MS methods, amounting to a total of 12 runs (see Materials and methods and Supplementary material for further details). The combined results of this approach ensured identification and relative quantitation of the different proteoforms of the six most abundant milk proteins in the samples, i.e., α_{S1}-, α_{S2}-, β- and κ-casein, α-lactalbumin and β-lactoglobulin, as presented in Fig. 2. Cumulatively, these proteins make up >90% of total protein in bovine milk (Brown et al., 2020; Dupont, Croguennec, Brodkorb, & Kouaouci, 2013), and thus also constitute the major source of protein glycation in bovine milk. Furthermore, within the current experimental setup, glycation was found to be the most abundant modification occurring during processing and storage of skim milk powder, with native and glycated proteoforms accounting for the greater part of the intact mass spectra (Supplementary material Fig. S1).

The thermised milk was the sample with the lowest intensity of processing in the current study. As expected, and in line with results shown in Fig. 1D, little to no glycation could be detected in this sample when analysed with the protein-centric approach (Fig. 2). Consequently, the thermised milk is representative of the endogenous repertoire of proteoforms for each of the six most abundant bovine milk proteins, including heterogeneity resulting from genetic polymorphism and varying levels of phosphorylation (Fig. 2). For simplicity, the O-glycosylated forms of κ-casein were not included in this analysis and the focus was kept on the more abundant non-glycosylated κ-casein proteoforms.

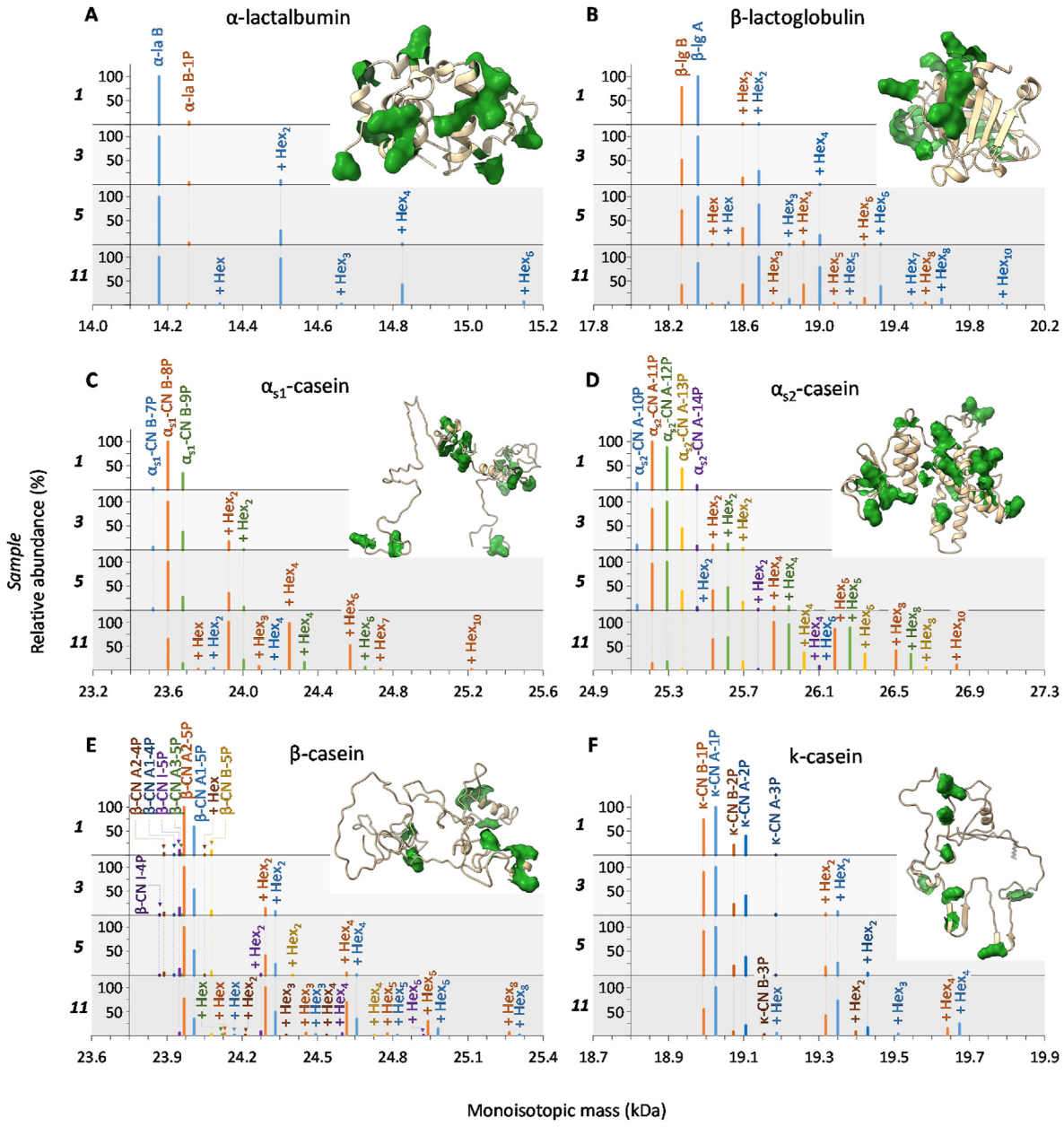

Fig. 2. Changes in the proteoform profiles of the six most abundant bovine milk proteins during production and storage of skim milk powder: α-lactalbumin (A), β-lactoglobulin (B), α_{S1}-casein (C), α_{S2}-casein (D), β-casein (E) and κ-casein (F). Shown are deconvoluted mass spectra depicting all proteoforms that were identified in thermised milk (1), preheated milk (3), fresh skim milk powder (5) and aged (159 d at 40 °C) skim milk powder (11). The nomenclature used for the different proteoforms is as proposed by the American Dairy Science Association Committee (Farrell Jr et al., 2004). The most intense proteoform in each spectrum was set to 100% abundance, relative to which the intensities of the other proteoforms were also expressed. The monomeric protein structures depicted in each panel were generated with ChimeraX 1.1; the lysine residues are highlighted in green. The whey protein structures are α-lactalbumin PDB 1F6S, chain B, and β-lactoglobulin PDB 1BEB, chain B. The casein structures are the energy-minimized monomers that were modelled into the submicellar casein particles as described by Farrell et al. (2013).

Preheating milk led to an increase in protein heterogeneity primarily due to lactosylation, detected as dihexose residue mass increments (+324 Da) at the intact protein level. Further increases in protein heterogeneity occurred particularly during skim milk powder storage, with increasing abundances of glycated proteoforms. Next to lactosylation, lower levels of glycation with monohexoses (+162 Da) were also observed on intact proteins, particularly in the aged powder.

The two proteins most susceptible to glycation were found to be α_{S2}-casein and β-lactoglobulin. The entire proteoform profile of the α_{S2}-casein shifted and diversified during production due to up to 2 glycation sites per protein molecule being occupied in the fresh powder (Fig. 2D). The non-glycated proteoforms still dominated the α_{S2}-casein proteoform profile of the fresh skim milk powder. This changed in the case of the aged powder, where up to 5 lactose residues per α_{S2}-casein molecule were detected. The α_{S2}-casein fraction in this sample became dominated by the proteoforms with 2 and 3 lactose residues per molecule.

In the case of β-lactoglobulin glycation could already be detected in the thermised milk, indicating that this protein is particularly susceptible to glycation even under very mild processing conditions (Fig. 2B). The degree of β-lactoglobulin glycation further

increased during processing resulting in up to 3 lactose residues per molecule detected in the fresh powder. The monolactosylated forms became dominant in the proteoform profiles along with the non-glycated forms. Considerable glycation occurred further during storage leading to up to 5 lactose residues per β-lactoglobulin molecule in the aged powder, with the proteoform profile shifting in favour of proteoforms carrying 0–2 lactose residues per molecule (Fig. 2B).

In the case of α_{S1}-casein, no glycation was detected in the thermised milk and only limited glycation could be detected during further processing with up to 1 lactose residue per molecule in both the preheated milk and in the fresh powder (Fig. 2C). In contrast, considerable glycation of the α_{S1}-casein occurred during storage of the powder, with up to 5 lactose residues per molecule being detected, whereby mono- and dilactosylated forms dominated the proteoform profile. The glycation of β-casein occurred gradually during processing, with monolactosylation of the proteins during the preheating of milk and dilactosylation occurring in the fresh powder (Fig. 2E). At the same time, the samples remained dominated by the non-glycated proteoforms. As also seen in the other proteins, considerable glycation of β-casein occurred during storage, yielding a proteoform profile dominated by the forms with 0–2 lactose residues, and a total of up to 4 lactose residues detected per molecule.

Unlike β-lactoglobulin, the other major whey protein, α-lactalbumin, proved to be far less susceptible to glycation (Fig. 2A). Limited glycation occurred during production with only monolactosylation detectable in the preheated milk and dilactosylation detectable in the fresh powder, but the samples remained dominated by the non-glycated proteoform of α-lactalbumin. The glycation advanced during storage with up to 3 lactose residues detectable per molecule and with the proteoform profile becoming dominated by the non-glycated and monolactosylated forms of α -lactalbumin (Fig. 2A). κ-Casein showed even lower susceptibility to glycation than α-lactalbumin (Fig. 2F). Monolactosylation could be detected in the preheated milk and fresh powder, and dilactosylation was detected in the aged powder. Irrespective of the sample, the proteoform profile of κ-casein remained dominated by the non-glycated proteoforms.

Irrespective of protein, no preference of glycation was observed towards specific genetic variants or phosphoproteoforms; instead, all forms of the same protein appeared to become similarly glycated (Supplementary material Table S2 and Fig. S2).

The percentage of glycated lysine in the sample calculated from the protein-centric mass spectrometry results was found to be in good accordance with the results of the blocked lysine analysis (Supplementary material Fig. S3). The percentage of glycated lysine estimated by mass spectrometry was 0.9 times that determined from the blocked lysine analysis. This difference could be caused by the assumptions made for calculation (i.e., glycation occurred exclusively on the six most abundant bovine milk proteins in the form of lysine lactosylation, excluding all other proteins and peptides, and other forms of glycation) and/or by the intensity suppression of glycation in the mass spectrometer due to reduction of the positive charges on the glycated residues (Fuerer et al., 2020).

3.3. Peptide-centric mass spectrometry

We complemented the data obtained by protein-centric mass spectrometry with peptide-centric mass spectrometry analyses to obtain a detailed site-specific characterisation of protein glycation. The four samples were digested with a series of proteases of different specificities (see Table 1), amounting to 28 LC-MS/MS runs, the cumulative results of which ensured high sequence coverage irrespective of glycation status (Supplementary material Fig. S4). Glycation was searched in the form of dihexose (i.e., lactosylation) and monohexose additions on lysine or arginine residues, or protein N-termini of the six most abundant milk proteins. Coverage of all putative glycation sites was achieved, except for the protein N-terminus of α_{S2}-casein. From this data it was possible to extract the major glycation sites and to quantify glycation site occupancy (see Materials and methods and Supplementary material for further details).

Fig. 3 provides site-specific information on glycation in the six most abundant milk proteins during production and storage of the skim milk powder. Data are shown only for the sites where glycation was detected at ≥1% of site occupancy. The vast majority of the detected glycation sites are lysine residues. To a lower extent, glycation was also detected at the protein N-terminus of α-lactalbumin (Fig. 3A). As also observed from the analysis at the protein level (Fig. 2), the peptide-centric approach revealed that glycation predominantly occurred in the form of lactosylation and, to a much lower extent, also in the form of glycation with monohexoses. Glycation with monohexoses was primarily detected in the aged powder. The results of the peptide-centric approach, and also in line with the results at the protein level (Fig. 2), show that β-lactoglobulin and α_{S2}-casein exhibited the highest susceptibility to glycation (Fig. 3B and D, respectively). For each of these proteins, up to 9 glycation sites were identified in the analysed samples, with one of them showing nearly 100% occupancy in the aged powder. Particularly residues Lys_{141} in β-lactoglobulin and Lys_{173} in α_{S2}-casein were detected at 37% and 100% occupancy already in the preheated milk. The occupancy of β-lactoglobulin Lys_{141} gradually increased to 95% glycation in the aged powder. As also observed from the protein-centric results, α_{S1}-casein showed lower susceptibility to glycation during milk powder production than α_{S2}-casein and β-lactoglobulin (<5% glycation site occupancy for all sites), but it developed considerable glycation during storage (Fig. 3C). Particularly α_{S1}-casein residues Lys_7, Lys_{42}, Lys_{83}, Lys_{102} and Lys_{103} had site occupancies in the range of 14–69% glycation in the aged powder. β-Casein showed lower susceptibility to glycation than the α_S-caseins (Fig. 3E), and the least susceptible proteins were found to be α-lactalbumin (Fig. 3A) and κ-casein (Fig. 3F), with only 4 and 3 glycation sites detected per protein molecule, respectively. For β-casein, κ-casein and α-lactalbumin, glycation site occupancies did not exceed 5% during the processing of skim milk into skim milk powder. Considerable increases in glycation site occupancy occurred during storage for β-casein residues Lys_{29}, Lys_{32}, Lys_{105}, and Lys_{169}, α-lactalbumin residues Glu_1, Lys_{62} and Lys_{98}, and κ-casein residue Lys_{24}, all increasing to the range of 10–27% occupancy. These findings are also supported by our data obtained using the protein-centric approach.

The overall picture emerging from the data shown in Figs. 2 and 3 is that little to no glycation was detected in the thermised milk. Low levels of glycation were induced by the heat treatment (60 s at 108 °C). Glycation increased in the skim milk powder as a result of the heat load during spray-drying. Considerable increases in glycation at existing and new sites occurred during the accelerated storage (159 d at 40 °C) of the skim milk powder. Glycation primarily occurred with dihexoses, i.e., lactose. Glycation with monohexoses was also detected to a lower extent, particularly in the aged skim milk powder. Moreover, specific lysine residues were substantially more prone to glycation than others, as further detailed below.

3.4. Glycated lysine motif analysis

The site-specific information obtained from the peptide-centric mass spectrometry approach (Fig. 3) indicated that not all putative glycation sites are equally susceptible to modification, with only a

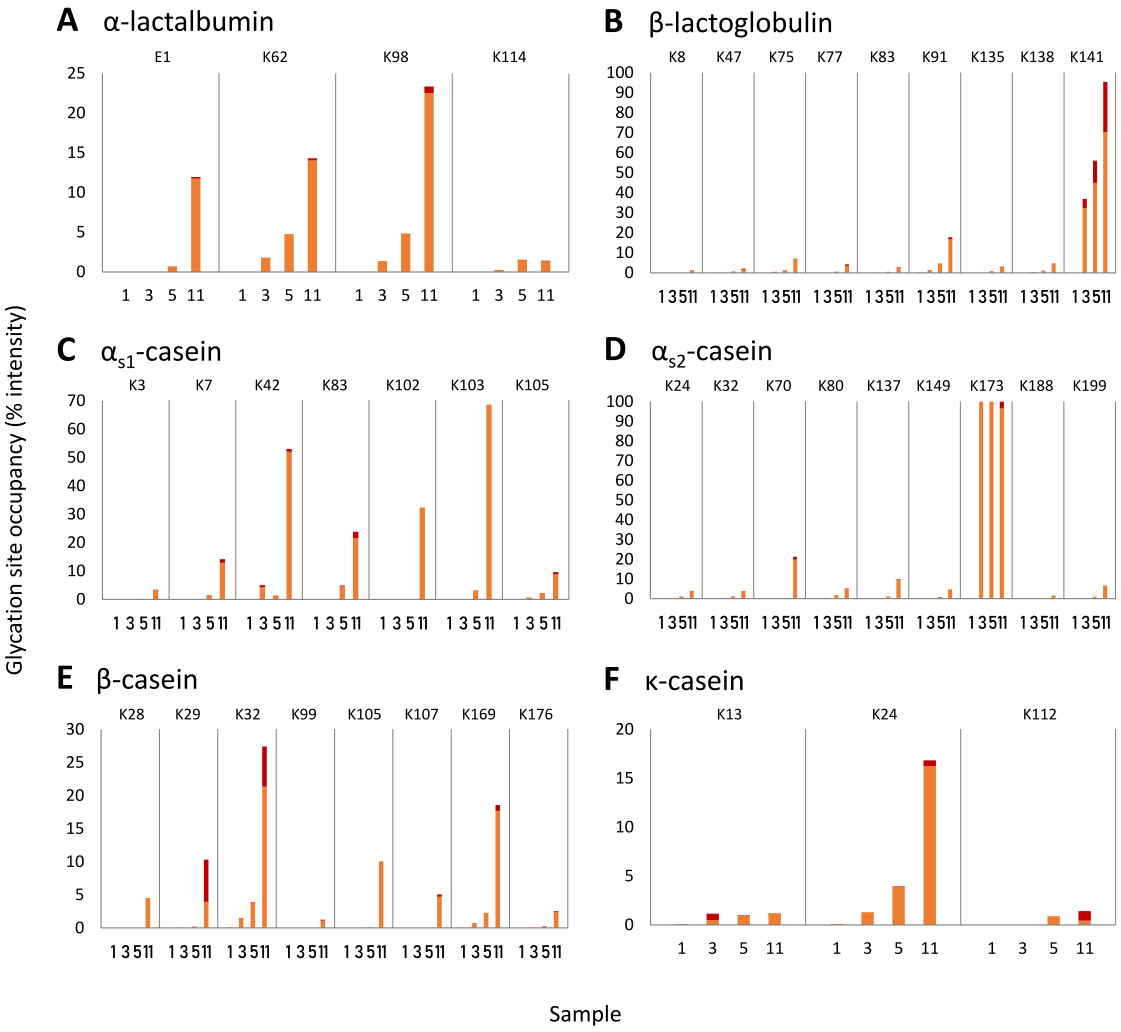

Fig. 3. Glycation hot-spots detected in the six most abundant bovine milk proteins. Glycation site occupancy (sum of intensities of peptides where a site has been identified as being glycated expressed as a percentage of total sum of intensities of all peptides containing the respective site) was determined by using a peptide-centric mass spectrometry approach, combining data obtained by using 7 protease combinations and a total of 28 LC-MS/MS runs. The glycation occupancy is shown for each of the four samples; thermised milk (sample 1), preheated milk (sample 3), fresh skim milk powder (sample 5) and aged (159 d at 40 °C) skim milk powder (sample 11). ■ indicates glycation through addition of a dihexose; ■ indicates glycation by a monohexose.

few selected sites in each protein representing so-called glycation hot-spots. Literature to date does not thoroughly describe the origin of these differences between the various lysine residues in milk proteins. In trying to address this question, we performed a sequence motif analysis of all the lysine residues in the six most abundant milk proteins to investigate whether the amino acid sequence containing the glycated or non-glycated residues correlates with their susceptibility towards glycation. The motif of the non-glycated lysine residues (63 residues) was subtracted from the motif of the glycated lysine residues with >5% site occupancy (22 residues) from the aged skim milk powder (159 d at 40 °C) to better highlight preferential sequence composition around the observed major glycation sites. This analysis, illustrated in Fig. 4 reveals for the first time that: (i) majorly glycated lysine residues in the most abundant bovine milk proteins were often preceded by a stretch of positively charged amino acids (blue), i.e., primarily lysine residues and to a smaller extent histidine residues, and (ii) at the C-terminal end of the glycated lysine a stretch of mostly aliphatic and hydrophobic amino acids, i.e., valine, leucine and isoleucine, proline and methionine was found to be enriched (Fig. 4, black).

4. Discussion

We carefully designed our study to fill knowledge gaps surrounding milk protein glycation in a complex and realistic dairy system that arise from sample selection, protease selection for peptide-centric mass spectrometry, general experimental design, and data analysis and interpretation. With this aim, we analysed a series of samples belonging to a single batch of bulk milk processed into skim milk powder and subsequently aged under controlled conditions. Assessment based on SDS-PAGE and the percentage of blocked lysine residues revealed glycation occurring during thermal processing and considerably increasing during storage of the powders (Fig. 1). Analysis of the four key samples, i.e., the thermised milk, the preheated milk, the fresh skim milk powder, and the aged skim milk powder (159 d at 40 °C), using a hybrid mass spectrometry approach combining protein-centric LC-MS and MS/MS (Fig. 2) with peptide-centric multi-protease LC-MS/MS (Fig. 3), yielded an in-depth view of glycation on the main bovine milk proteins. The protein-centric mass spectrometry results gave a quantitative overview of the different proteoforms and of their extent of glycation, while the peptide-centric approach

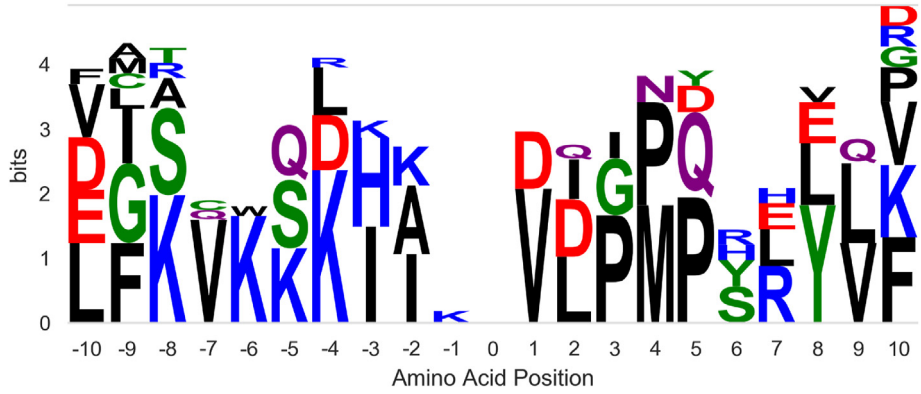

Fig. 4. Enriched amino acid sequence motif surrounding glycated lysine residues in bovine milk proteins. Position probability matrix calculated based on the subtraction of position frequency matrices of the 63 non-glycated lysine residues (<5% glycation site occupancy) from the 22 majorly glycated lysine residues (≥5% glycation site occupancy) in the six most abundant milk proteins in the aged (159 d at 40 °C) skim milk powder. Position 0 indicates the glycated lysine residues; positions −10 to −1 and 1 to 10 indicate the 10 amino acids N-terminal and C-terminal of the glycated lysine in the amino acid sequence, respectively.

complemented the aforementioned results with qualitative information on the localisation of the glycation hot-spots.

We clearly observed that not all lysine residues are equally targeted by glycation. Among the glycation hot-spots we identified, several were already known, but novel ones gain visibility thanks to our experimental design. Analysing milk protein glycation solely in a tryptic digest is biased against identifying glycation occurring in lysine and arginine-rich regions, such as residues Lys_{135}, Lys_{138} and Lys_{141} in β-lactoglobulin, Lys_{102}, Lys_{103} and Lys_{105} in α_{S1}-casein, Lys_{21} and Lys_{24} in α_{S2}-casein, Lys_{28}, Lys_{29} and Lys32 in β-casein or Lys_{111}, Lys_{112} and Lys_{116} in κ-casein. While only poor coverage of these sites is typically achieved relying on trypsin hydrolysis (Milkovska-Stamenova & Hoffmann, 2016a,b), our multi-protease approach revealed these protein regions to harbour several glycation hot-spots, as depicted in Fig. 3. Conversely, glycation hot-spots located on tryptic macropeptides that are outside the range of conventional peptide-centric LC-MS/MS methods, are often missed as well. This is the case of residues Lys_{70} and Lys_{169} that we here reveal to be two of the main glycation sites in α_{S2}-casein and β-casein, respectively. Lys_{70} is located on the α_{S2}-casein tryptic 46–70; its glycated form results in a missed cleavage at Lys_{70} and ends up on the longer tryptic peptide 46–76 of 31 residues in length. Lys_{169} of β-casein is located on an even longer tryptic macropeptide of residues 114–169 (56 amino acids); glycation and missed cleavage at Lys_{169} puts it on the longer tryptic peptide 114–176 of 63 amino acid residues in length. While these glycation hot-spots were also missed in trypsin-based peptide-centric approaches (Milkovska-Stamenova & Hoffmann, 2016a,b), we have shed light on them here on account of the high sequence coverage (Supplementary material Fig. S4) obtained with our multi-protease approach.

Identification and relative quantitation of the glycation hot-spots that develop during production and storage of skim milk powder still leaves the question unanswered as to why specific sites are favoured over others. At the primary sequence level, we reveal in this study, for the first time, a preferred motif where the glycation hot-spot in milk proteins is preceded to the N-terminal by positively charged amino-acids (lysine and histidine residues), and to the C-terminal it is followed mostly by aliphatic and hydrophobic amino acids (valine, leucine and isoleucine, proline and methionine residues; see Fig. 4). The study of Johansen, Kiemer, and Brunak (2006) was one of the first to analyse a putative glycation motif in a combined dataset of 20 proteins, comprising 89 glycated and 126 non-glycated lysine residues, originating from in vivo and in vitro mammalian studies. Despite considerable differences in the study systems and conditions, their findings indicate the occurrence of basic residues to the N-terminal and of acidic residues to the C-terminal of the glycated lysine, in agreement with our data. Furthermore, Johansen et al. (2006) argued that basic and acidic residues located in the proximity of the glycation hot-spot, either on the primary sequence or on the 3D structure, act as catalysts for the Amadori rearrangement. A much broader study in a similar direction from Zhang et al. (2011) investigated the glycation motif occurring in over 3700 blood plasma and erythrocyte proteins due to diabetes. Although less pronounced, their primary sequence motif also supports the enrichment of basic residues N-terminal, and of aliphatic and acidic residues C-terminal of glycation hot-spots, as also observed in our study (Fig. 4).

Next to revealing the preferential motif in the primary sequence of the milk proteins, we considered whether higher-order structural features may also correlate to the glycation hot-spots revealed in our study. The hollow β-barrel structure with a hydrophobic pocket of β-lactoglobulin (PDB entry 1BEB; Brownlow et al., 1997) places most of the hydrophilic residues at the surface of the molecule, rendering the majority of the lysine residues accessible for glycation in the globular protein. However, as depicted in Fig. 1B, most of the β-lactoglobulin no longer entered the non-reducing gel following preheating, indicating that the protein unfolded and likely formed disulphide-linked aggregates.

Denaturation and unfolding of the molecule can further expose reactive lysine residues, increasing their accessibility for lactosylation, explaining the high degree of glycation that was observed for β-lactoglobulin (Figs. 2B and 3B). While also a globular protein, the structure of α-lactalbumin is more compact and stabilised by a Ca^{2+} ion and 4 disulphide bridges (Permyakov & Berliner, 2000), which renders it resistant to both glycation (Figs. 2A and 3A) and proteolytic degradation. As observed from the SDS-polyacrylamide gels in Fig. 1B and C, most of the α-lactalbumin remained presumably still compactly folded following preheating of the milk, possibly explaining its observed low level of glycation. The glycation hot-spots detected in α-lactalbumin (Fig. 3A), i.e., residues Glu_1, Lys_{62} and Lys_{98}, and to a lesser extent Lys_{114}, all have surface-exposed side-chains that are not involved in the secondary structure of the protein (PDB entry 1F6S; Chrysina, Brew, & Acharya, 2000). The α_S-caseins underwent the highest levels of glycation (Fig. 2C and D and 3C,D) despite their localisation in the core of the casein micelles (Huppertz et al., 2017). This can possibly be attributed to the scarcity of casein secondary structure (Farrell, Brown, & Malin, 2013) combined with the open and porous structure of the casein micelles (Bouchoux, Gesan-Guiziou, Pérez, & Cabane, 2010), allowing accessibility of lactose into the micelle.

Extensive research into casein structure has resulted in energy-minimised modelled structures for homo-oligomers and sub-micellar casein particles (Farrell, Brown, Hoagland, & Malin, 2003), relative to which the glycation hot-spots can be considered. Glycation of α_{S1}-casein was only detected in the N-terminal part of the protein, up until Lys_{105} (Fig. 3C). The C-terminal part of the protein is for the most part involved in intermolecular interactions (Farrell Jr, Malin, Brown, & Mora-Gutierrez, 2009; Kumosinski, Brown, & Farrell Jr, 1994a; Kumosinski, King, & Farrell, 1994b). The resulting limited solvent and lactose accessibility may explain the absence of glycation observed in this region of the protein. The lysine residues in κ-casein are found for the most part in the para-κ-casein region, which is also the region of interaction with the casein micelle, limiting putative glycation of the protein. Lys_{24} has high solvent exposure (Kumosinski et al., 1994b) and it was detected as one of the sites to undergo the highest level of glycation (Fig. 3F). β-Casein is the most proline-rich of the caseins (35 proline residues out of 209 amino acids), with therefore the least secondary structure, the most open conformation and the highest lactose accessibility. Glycation sites Lys_{28}, Lys_{29} and Lys_{32}, and Lys_{99}, Lys_{105} and Lys_{107} (Fig. 3E) are clustered close to each other and are found in 2 lobes of the protein localised polar opposite from the dimerisation region (Kumosinski, Brown, & Farrell Jr, 1993). All of the sites mentioned above are solvent-exposed. Furthermore, Lys_{28}-Lys_{29}, Lys_{105}-His_{106} and Lys_{107}-Glu_{108} are the major plasmin hydrolysis sites in β-casein (Eigel et al., 1984), they are easily accessible by this large enzyme, and consequently also accessible by lactose, and prone to glycation. Glycation sites Lys_{169} and Lys_{176} (Fig. 3E) are in close proximity of the C-terminal tail of the protein involved in intermolecular interactions (Kumosinski et al., 1994a). Residues Lys_{169} and Lys_{176} are solvent-exposed within a large open cavity in the β-casein molecule, allowing accessibility of lactose molecules.

In light of these findings, rather than an absolute affinity of the reducing carbohydrate molecules to the amino acid sequence, we believe the sequence motif to be correlated with protein conformation. Thereby, lysine residues found in this motif are likely to be surface-exposed in the structure of the proteins, and consequently accessible for glycation.

5. Conclusions

Our study provides new insights into the development, advancement and localisation of glycation on the milk proteins. The sample composition and the processing and storage conditions can be linked directly to the development and advancement of glycation, highlighting the importance of studying representative samples. Glycation was triggered and accelerated during the processing steps where a high heat load was brought to the samples, and considerably increased during storage of the skim milk powder. Glycation hot-spot localisation and comparison of putative glycation sites can only be achieved with careful experimental design that ensures high protein sequence coverage. We found here protein-specific susceptibility towards glycation and the glycation hot-spots were revealed to have a preferential motif on the primary sequence, and are surface-exposed and solvent-accessible on the higher-order protein structures. Better understanding of glycation is prerequisite for limiting the extent of glycation during production and storage, and thereby achieving high quality dairy ingredients and products.

Declaration of competing interest

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

Acknowledgements

We acknowledge support from the Dutch Research Council (NWO) in the framework of the Innovation Fund for Chemistry (SATIN project 731.017.202). We thank Max Hoek (Biomolecular Mass Spectrometry and Proteomics, Bijvoet Centre for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, Utrecht University; Netherlands Proteomics Centre) for assistance with the analysis of the glycation motif in the primary sequence of the proteins. We thank Kelly A. Dingess (Biomolecular Mass Spectrometry and Proteomics, Bijvoet Centre for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, Utrecht University; Netherlands Proteomics Centre) for proofreading the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.idairyj.2022.105340.

References

Aalaei, K., Rayner, M., & Sjöholm, I. (2019). Chemical methods and techniques to monitor early Maillard reaction in milk products; A review. *Critical Reviews in Food Science and Nutrition, 59*, 1829–1839.

Arena, S., Renzone, G., D'Ambrosio, C., Salzano, A. M., & Scaloni, A. (2017). Dairy products and the maillard reaction: A promising future for extensive food characterization by integrated proteomics studies. *Food Chemistry, 219*, 477–489.

Baye, E., de Courten, M. P. J., Walker, K., Ranasinha, S., Earnest, A., Forbes, J. M., et al. (2017). Effect of dietary advanced glycation end products on inflammation and cardiovascular risks in healthy overweight adults: A randomised crossover trial. *Scientific Reports, 7*, Article 4123.

Birlouez-Aragon, I., Pischetsrieder, M., Leclere, J., Morales, F. J., Hasenkopf, K., Kientsch-Engel, R., et al. (2004). Assessment of protein glycation markers in infant formulas. *Food Chemistry, 87*, 253–259.

van Boekel, M. (1998). Effect of heating on Maillard reactions in milk. *Food Chemistry, 62*, 403–414.

Bondt, A., Hoek, M., Tamara, S., de Graaf, B., Peng, W., Schulte, D., et al. (2021). Human plasma IgG1 repertoires are simple, unique, and dynamic. *Cell Systems, 12*, 1131–1143.e5.

Boor, K. J., Wiedmann, M., Murphy, S., & Alcaine, S. (2017). A 100-year review: Microbiology and safety of milk handling. *Journal of Dairy Science, 100*, 9933–9951.

Bouchoux, A., Gesan-Guiziou, G., Pérez, J., & Cabane, B. (2010). How to squeeze a sponge: Casein micelles under osmotic stress, a SAXS study. *Biophysical Journal, 99*, 3754–3762.

Brownlow, S., Cabral, J. H. M., Cooper, R., Flower, D. R., Yewdall, S. J., Polikarpov, I., et al. (1997). Bovine β-lactoglobulin at 1.8 Å resolution—still an enigmatic lipocalin. *Structure, 5*, 481–495.

Brown, B. A., Zeng, X., Todd, A. R., Barnes, L. F., Winstone, J. M. A., Trinidad, J. C., et al. (2020). Charge detection mass spectrometry measurements of exosomes and other extracellular particles enriched from bovine milk. *Analytical Chemistry, 92*, 3285–3292.

Cardoso, H. B., Wierenga, P. A., Gruppen, H., & Schols, H. A. (2018). Maillard induced glycation behaviour of individual milk proteins. *Food Chemistry, 252*, 311–317.

Carulli, S., Calvano, C. D., Palmisano, F., & Pischetsrieder, M. (2011). MALDI-TOF MS characterization of glycation products of whey proteins in a glucose/galactose model system and lactose-free milk. *Journal of Agricultural and Food Chemistry, 59*, 1793–1803.

Cattaneo, S., Stuknytė, M., Masotti, F., & De Noni, I. (2017). Protein breakdown and release of β-casomorphins during in vitro gastro-intestinal digestion of sterilised model systems of liquid infant formula. *Food Chemistry, 217*, 476–482.

Chevalier, F., Chobert, J. M., Dalgalarrondo, M., Choiset, Y., & Haertlé, T. (2002). Maillard glycation of β-lactoglobulin induces conformation changes. *Nahrung, 46*, 58–63.

Chrysina, E. D., Brew, K., & Acharya, K. R. (2000). Crystal structures of apo- and holobovine α-lactalbumin at 2.2-Å resolution reveal an effect of calcium on inter-lobe interactions. *Journal of Biological Chemistry, 275*, 37021–37029.

Delatour, T., Hegele, J., Parisod, V., Richoz, J., Maurer, S., Steven, M., et al. (2009). Analysis of advanced glycation endproducts in dairy products by isotope dilution liquid chromatography–electrospray tandem mass spectrometry. The particular case of carboxymethyllysine. *Journal of Chromatography A, 1216*, 2371–2381.

Delgado, T., Corzo, N., Santa-María, G., Jimeno, M., & Olano, A. (1992). Determination of furosine in milk samples by ion-pair reversed phase liquid chromatography. *Chromatographia, 33*, 374–376.

Deng, Y., Wierenga, P. A., Schols, H. A., Sforza, S., & Gruppen, H. (2017). Effect of Maillard induced glycation on protein hydrolysis by lysine/arginine and non-lysine/arginine specific proteases. *Food Hydrocolloids, 69*, 210–219.

Dupont, D., Croguennec, T., Brodkorb, A., & Kouaouci, R. (2013). Quantitation of proteins in milk and milk products. In P. L. H. McSweeney, & P. F. Fox (Eds.), *Advanced dairy chemistry* (4th edn., Vol. 1A, pp. 87–134). Boston, MA, USA: Springer.

Eigel, W. N., Butler, J. E., Ernstrom, C. A., Farrell, Jr., H. M., Harwalkar, V. R., Jenness, R., et al. (1984). Nomenclature of proteins of cow's milk: Fifth revision. *Journal of Dairy Science, 67*, 1599–1631.

Farrell, H. M., Brown, E. M., Hoagland, P. D., & Malin, E. L. (2003). Higher order structures of the caseins: A paradox?. In P. F. Fox, & P. L. H. McSweeney (Eds.), *Advanced dairy chemistry—1 Proteins* (3rd edn., Vol. IA, pp. 203–231). Boston, MA, USA: Springer.

Farrell, H. M., Brown, E. M., & Malin, E. L. (2013). Higher order structures of the caseins: A paradox?. In P. L. H. McSweeney, & P. F. Fox (Eds.), *Advanced dairy chemistry* (4th edn., Vol. IA, pp. 161–184). Boston, MA, USA: Springer.

Farrell, Jr., H. M., Jimenez-Flores, R., Bleck, G. T., Brown, E. M., Butler, J. E., Creamer, L. K., Hicks, C. L., Hollar, C. M., Ng-Kwai-Hang, K. F., & Swaisgood, H. E. (2004). Nomenclature of the proteins of cows' milk –Sixth revision. *Journal of Dairy Science, 87*(6), 1641–1674.

Farrell, Jr., H. M., Malin, E. L., Brown, E. M., & Mora-Gutierrez, A. (2009). Review of the chemistry of α_{S2}-casein and the generation of a homologous molecular model to explain its properties. *Journal of Dairy Science, 92*, 1338–1353.

Franc, V., Zhu, J., & Heck, A. J. R. (2018). Comprehensive proteoform characterization of plasma complement component C8$\alpha\beta\gamma$ by hybrid mass spectrometry approaches. *Journal of the American Society for Mass Spectrometry, 29*, 1099–1110.

Fuerer, C., Jenni, R., Cardinaux, L., Andetsion, F., Wagnière, S., Moulin, J., et al. (2020). Protein fingerprinting and quantification of β-casein variants by ultra-performance liquid chromatography–high-resolution mass spectrometry. *Journal of Dairy Science, 103*, 1193–1207.

Gómez-Narváez, F., Perez-Martinez, L., & Contreras-Calderon, J. (2019). Usefulness of some Maillard reaction indicators for monitoring the heat damage of whey powder under conditions applicable to spray drying. *International Dairy Journal, 99*, Article 104553.

Gonzales, A. P., Naranjo, G., Leiva, G., & Malec, L. (2010). Maillard reaction kinetics in milk powder: Effect of water activity at mild temperatures. *International Dairy Journal, 20*, 40–45.

Hegele, J., Buetler, T., & Delatour, T. (2008). Comparative LC–MS/MS profiling of free and protein-bound early and advanced glycation-induced lysine modifications in dairy products. *Analytica Chimica Acta, 617*, 85–96.

Hertz, G. Z., & Stormo, G. D. (1999). Identifying DNA and protein patterns with statistically significant alignments of multiple sequences. *Bioinformatics, 15*, 563–577.

Huppertz, T., Gazi, I., Luyten, H., Nieuwenhuijse, H., Alting, A., & Schokker, E. (2017). Hydration of casein micelles and caseinates: Implications for casein micelle structure. *International Dairy Journal, 74*, 1–11.

Johansen, M. B., Kiemer, L., & Brunak, S. (2006). Analysis and prediction of mammalian protein glycation. *Glycobiology, 16*, 844–853.

Krause, R., Knoll, K., & Henle, T. (2003). Studies on the formation of furosine and pyridosine during acid hydrolysis of different Amadori products of lysine. *European Food Research and Technology, 216*, 277–283.

Kumosinski, T. F., Brown, E. M., & Farrell, Jr., H. M. (1993). Three-dimensional molecular modeling of bovine caseins: An energy-minimized β-casein structure. *Journal of Dairy Science, 76*, 931–945.

Kumosinski, T. F., Brown, E. M., & Farrell, Jr., H. M. (1994a). Predicted energy-minimized α_{S1}-casein working model. In T. F. Kumosinski, & M. N. Liebman (Eds.), *Molecular modeling, from virtual tools to real problems* (pp. 368–391). Washington, DC, USA: ACS Symposium Series.

Kumosinski, T. F., King, G., & Farrell, H. M. (1994b). An energy-minimized casein submicelle working model. *Journal of Protein Chemistry, 13*, 681–700.

Leiva, G. E., Naranjo, G. B., & Malec, L. S. (2017). A study of different indicators of Maillard reaction with whey proteins and different carbohydrates under adverse storage conditions. *Food Chemistry, 215*, 410–416.

van Lieshout, G. A. A., Lambers, T. T., Bragt, M. C. E., & Hettinga, K. A. (2020). How processing may affect milk protein digestion and overall physiological outcomes: A systematic review. *Critical Reviews in Food Science and Nutrition, 60*, 2422–2445.

Milkovska-Stamenova, S., & Hoffmann, R. (2016a). Hexose-derived glycation sites in processed bovine milk. *Journal of Proteomics, 134*, 102–111.

Milkovska-Stamenova, S., & Hoffmann, R. (2016b). Identification and quantification of bovine protein lactosylation sites in different milk products. *Journal of Proteomics, 134*, 112–126.

Nyakayiru, J., van Lieshout, G. A. A., Trommelen, J., van Kranenburg, J., Verdijk, L. B., Bragt, M. C. E., et al. (2020). The glycation level of milk protein strongly modulates post-prandial lysine availability in humans. *British Journal of Nutrition, 123*, 545–552.

Oliver, C. M. (2011). Insight into the glycation of milk proteins: An ESI-and MALDI-MS perspective. *Critical Reviews in Food Science and Nutrition, 51*, 410–431.

Olsen, J. V., Ong, S.-E., & Mann, M. (2004). Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Molecular & Cellular Proteomics, 3*, 608–614.

Permyakov, E. A., & Berliner, L. J. (2000). α-Lactalbumin: Structure and function. *FEBS Letters, 473*, 269–274.

Renzone, G., Arena, S., & Scaloni, A. (2015). Proteomic characterization of intermediate and advanced glycation end-products in commercial milk samples. *Journal of Proteomics, 117*, 12–23.

Schuck, P., Blanchard, E., Dolivet, A., Méjean, S., Onillon, E., & Jeantet, R. (2005). Water activity and glass transition in dairy ingredients. *Lait, 85*, 295–304.

Sharma, S. D., & Barone, M. (2019). Deleterious consequences of dietary advanced glycation end products on human health due to oxidative stress and inflammation. In *Dietary patterns, food chemistry and human health* (pp. 1–13). Boston, MA, USA: Springer.

Siciliano, R. A., Mazzeo, M. F., Arena, S., Renzone, G., & Scaloni, A. (2013). Mass spectrometry for the analysis of protein lactosylation in milk products. *Food Research International, 54*, 988–1000.

Tamara, S., Franc, V., & Heck, A. J. R. (2020). A wealth of genotype-specific proteoforms fine-tunes hemoglobin scavenging by haptoglobin. *Proceedings of the National Academy of Sciences, 117*, 15554–15564.

Teodorowicz, M., Van Neerven, J., & Savelkoul, H. (2017). Food processing: The influence of the maillard reaction on immunogenicity and allergenicity of food proteins. *Nutrients, 9*, Article 835.

Toda, M., Hellwig, M., Henle, T., & Vieths, S. (2019). Influence of the Maillard reaction on the allergenicity of food proteins and the development of allergic inflammation. *Current Allergy and Asthma Reports, 19*, Article 4.

Tsiatsiani, L., & Heck, A. J. R. (2015). Proteomics beyond trypsin. *FEBS Journal, 282*, 2612–2626.

van de Waterbeemd, M., Tamara, S., Fort, K. L., Damoc, E., Franc, V., Bieri, P., et al. (2018). Dissecting ribosomal particles throughout the kingdoms of life using advanced hybrid mass spectrometry methods. *Nature Communications, 9*, Article 2493.

Yang, Y., Liu, F., Franc, V., Halim, L. A., Schellekens, H., & Heck, A. J. R. (2016). Hybrid mass spectrometry approaches in glycoprotein analysis and their usage in scoring biosimilarity. *Nature Communications, 7*, Article 13397.

Zhang, Q., Monroe, M. E., Schepmoes, A. A., Clauss, T. R. W., Gritsenko, M. A., Meng, D., et al. (2011). Comprehensive identification of glycated peptides and their glycation motifs in plasma and erythrocytes of control and diabetic subjects. *Journal of Proteome Research, 10*, 3076–3088.

Zhu, J., Lin, Y.-H., Dingess, K. A., Mank, M., Stahl, B., & Heck, A. J. R. (2020). Quantitative longitudinal inventory of the N-glycoproteome of human milk from a single donor reveals the highly variable repertoire and dynamic site-specific changes. *Journal of Proteome Research, 19*, 1941–1952.