PROTEOMIC CHARACTERIZATION OF INTERMEDIATE AND ADVANCED GLYCATION END-PRODUCTS IN COMMERCIAL MILK SAMPLES

Giovanni Renzone, Simona Arena and Andrea Scaloni

Proteomics & Mass Spectrometry Laboratory, ISPAAM, National Research Council, Naples, Italy

Keywords: milk; thermal treatment; advanced-glycation end-products; lactosylation

Correspondence to:

Andrea Scaloni, PhD

Proteomics & Mass Spectrometry Laboratory

ISPAAM, National Research Council

via Argine 1085, 80147 Naples, Italy

E-mail: andrea.scaloni@ispaam.cnr.it

Tel: +39 0815966006; Fax: +39 0815965291

Abstract

The Maillard reaction consists of a number of chemical processes affecting the structure of the proteins present in foods. We previously accomplished the proteomic characterization of the lactosylation targets in commercial milk samples. Although characterizing the early modification derivatives, this analysis did not describe the corresponding advanced glycation end-products (AGEs), which may be formed from the further oxidation of former ones or by reaction of oxidized sugars with proteins, when high temperatures are exploited. To fill this gap, we have used combined proteomic procedures for the systematic characterization of the lactosylated and AGEs-containing proteins from the soluble and milk fat globule membrane fraction of various milk products. Beside to confirm all lactulosyl-lysines described previously, 40 novel lactosylation sites were identified. More importantly, 308 additional intermediate and advanced glycoxidation derivatives (including cross-linking adducts) were characterized in 31 proteins, providing the widest qualitative inventory of modified species ascertained in commercial milk samples so far. Amadori adducts with glucose/galactose, their dehydration products, carboxymethyl-lysine, and glyoxal-, 3-deoxyglucosone/3-deoxygalactosone and 3-deoxylactosone-derived dihydroxyimidazolines and/or hemiaminals were the most frequent derivatives observed. Depending on thermal treatment, a variable number of modification sites was identified within each protein; their number increased with harder food processing conditions. Among the modified proteins, species involved in assisting the delivery of nutrients, defense response against pathogens and cellular proliferation/differentiation were highly affected by AGEs formation. This may lead to a progressive decrease of the milk nutritional value, as it reduces the protein functional properties, abates the bioavailability of the essential amino acids and eventually affects food digestibility. These aspects are of particular importance in products intended for infant diet, such as milk powders and infant formulas.

Abbreviations

AGEs, advanced glycation end products; G, glyoxal; MG, methylglyoxal; 3-DG, 3-deoxyglucosone; 3-DGal, 3-deoxygalactosone; 3-DLact, 3-deoxylactosone; 3-DPen, 3-deoxypentosone; Trios, triosone; Lact, the Amadori compound with lactose; Lact-H_2O, the dehydrated Amadori compound with lactose; Lact-$2H_2O$, the bis-dehydrated Amadori compound with lactose; Hex, the Amadori compound with lactose glucose/galactose; Hex-H_2O, the dehydrated Amadori compound with lactose glucose/galactose; Hex-$2H_2O$, the bis-dehydrated Amadori compound with lactose glucose/galactose; CML, carboxymethyllysine; CEL, carboxyethyllysine; PYR, pyrraline; RPYR, Arg-pyrimidine; PENT, pentosidine; DOLD, 3-DG/3-DGal-derived lysine dimer; DOGDIC, 3-DG/3-DGal-derived imidazolium cross-link; G-DH, G-derived dihydroxyimidazoline; MG-DH, MG-derived dihydroxyimidazoline; Trios-DH, Trios-derived dihydroxyimidazoline; 3-DPen-DH, 3-DPen-derived dihydroxyimidazoline; 3-DG-DH, 3-DG-derived dihydroxyimidazoline; 3-DGal-DH, 3-DGal-derived dihydroxyimidazoline; 3-DLact-DH, 3-DLact-derived dihydroxyimidazoline; G-H, G-derived hydroimidazolone; MG-H, MG-derived hydroimidazolone; 3-DG-H, 3-DG-derived hydroimidazolone; 3-DGal-H, 3-DGal-derived hydroimidazolone; 3-DLact-H, 3-DLact-derived hydroimidazolone; 3-DPen-H, 3-DPen-derived hydroimidazolone; G-He, G-derived hemiaminal; MG-He, MG-derived hemiaminal; Trios-He, Trios-derived hemiaminal; 3-DPen-He, 3-DPen-derived hemiaminal; 3-DG-He, 3-DG-derived hemiaminal; 3-DGal-He, 3-DGal-derived hemiaminal; 3-DLact-He, 3-DLact-derived hemiaminal; Lact-He, lactosone-derived hemiaminal; G-SB, G-derived Schiff base; MG-SB, MG-derived Schiff base.

Introduction

In order to preserve microbiological safety and to prolong shelf life, milk products are submitted to thermal treatment prior to retail and consumption. These processes induce significant physicochemical changes in milk proteins, depending on the duration/extent of the heating procedure. The most thoroughly studied thermal modification of milk proteins is the Maillard reaction, a non-enzymatic glycation in which the carbonyl group of reducing sugars (primarily lactose) reacts with the N^{ε}-group of Lys residues to yield the Amadori product lactulosyl-lysine (Lact) [1-3]. Reactivity at protein N-terminus is also observed. When higher heat impact is applied during milk processing, Lact is degraded and further converted into advanced glycation end products (AGEs) [2,3]. Oxidative degradation product N^{ε}-carboxymethyllysine (CML) was identified as the major AGE formed in heated milk models [2-4]. Additional AGEs recognized in processed milk include N^{ε}-carboxyethyllysine (CEL), pyrraline (PYR), oxalic acid monolysinylamide and pentosidine (PENT) [2,3,5]. In parallel, AGEs can also be generated by the reaction of protein Arg or Lys with sugar (lactose) heating degradation products, i.e. glyoxal (G), methylglyoxal (MG), 3-deoxyglucosone (3-DG), 3-deoxygalactosone (3-DGal), 3-deoxypentosone (3-DPen) (Supplementary Fig. S1) [6-9]. When model peptides/proteins were treated with these dicarbonyl compounds, the formation of the AGEs mentioned above and of : i) G-, MG-, 3-DG- and 3-DGal-derived hemiaminals (G-He, MG-He, 3-DG-He, and 3-DGal-He), ii) G-, MG-, 3-DG- and 3-DGal-derived dihydroxyimidazolines (G-DH, MG-DH, 3-DG-DH, and 3-DGal-DH), iii) G-, MG-, 3-DG-, and 3-DGal-derived hydroimidazolones (G-H, MG-H, 3-DG-H, and 3-DGal-H), iv) di- and tetra-hydropyrimidines (DHP and THP) was ascertained (Supplementary Fig. S2) [9,10]. Lipid peroxidation products, by generating intermediate dicarbonyl compounds, can also contribute to the production of specific AGEs, such as CML [9,11,12].

Depending on their relative stability, AGEs were initially identified by dedicated immunoassays or amino acid analysis on exhaustive molecular hydrolysates, missing information

on the nature of the modified proteins and the corresponding modification site(s) [2,3,5,9]. In the last decade, various MS approaches have been developed for the detection and the direct assignment of AGEs in proteins from biological matrices and foods [9]. Regarding milk proteins, peptide mapping experiments were performed on isolated model components heated *in vitro* with various sugars to simulate food processing conditions or on a very reduced number of proteins isolated by chromatographic or electrophoretic procedures from commercial milk samples [1,4,9,13-18 and references therein]. Although a number of studies was realized in this context, AGEs assignment was generally restricted to the case of CML adducts, complementing data on oxidized Lys, Met, Cys, Tyr and Trp derivatives, lysinoalanine, histidinoalanine, pyroglutamate and deamidated Asn [2-4,13,19]. Recently, untargeted LC-ESI-MS/MS multiple reaction monitoring (MRM) procedures were used to evaluate the different AGEs present in β-lactoglobulin (β-LG) from heated milk models [19]. When a derived targeted MRM procedure was applied to the analysis of β-LG from commercial milk samples, only Lact, CML, CEL, MG-H, MG-DH and 3-DG-H adducts were detected and assigned to specific residues [19,20]. Some of these derivatives were already identified in β-casein treated *in vitro* with glucose or MG [21]. Although highly representative of the main AGEs (mainly CML) present in the most abundant milk proteins, all studies mentioned above missed the global picture deriving from the application of shotgun proteomic approaches to milk products, which may allow the simultaneous assignment of different glyco-oxidative modifications to the number of components present in this food.

In the latter context, we recently used integrated proteomic procedures for the global characterization of the various lactosylation protein targets in different commercial milk products [22,23]. Soluble milk proteins enriched for less abundant components by the ProteoMiner technology and protein components from purified milk fat globule (MFG) particles were resolved by SDS-PAGE, enzymatically digested, and enriched for lactosylated peptides by affinity chromatography on *m*-aminophenylboronic acid-derivatized agarose. Shotgun characterization of

the lactosylated peptides was obtained by nLC–ESI–LIT–MS/MS, using electron transfer dissociation (ETD) fragmentation combined with supplemental collisional activation. We and others verified that CID fragmentation of glycated peptides generates mass spectra characterized by non-sequence-informative ion species associated with the cleavage of the glycosidic bond, and rare b- and y-ions [22,24]. When applied to the soluble milk fraction, this strategy allowed the identification of 271 non-redundant lactosylation sites in 33 proteins [22]. Similarly, 157 novel non-redundant lactosylation sites were identified in 35 MFG proteins [23], in addition to the 153 ones present in other 21 soluble proteins adsorbed on MFG membranes. Proteins involved in nutrient delivery, having a defensive function against bacteria/viruses or promoting cell proliferation/differentiation were highly affected by lactosylation. Shotgun analysis characterized a variable number of lactosylation sites within each protein; their number increased with harder milk processing conditions. This may lead to a progressive decrease of the milk nutritional value, as it reduces the bioavailability of essential amino acids, and may affect function, digestibility and allergenic potential of individual proteins [14,23,25,26].

Taking advantage of our previous experience on lactosylated species, a similar shotgun approach was applied to the global assignment of the AGEs present in proteins from raw and commercial milk samples. The results presented here provide the widest qualitative inventory of assigned intermediate and advanced glyco-oxidation protein adducts in milk products ascertained so far, providing important dietary information since AGEs-containing milk proteins are ingested daily from humans and, as unique polypeptide source, from infants.

Materials and Methods

Materials and sample preparation

Three samples of raw, pasteurized, ultra-high temperature (UHT), and powdered for infant nutrition milk (800 ml) were obtained from local farmers or different commercial sources. While

identical results (± 5%) were obtained for all samples of raw, pasteurized or UHT milk, reconstituted powdered milk samples (14% w/v) showed protein, carbohydrate, lipid and mineral content values ranging in 13.1-16.6, 79.2-84.8, 22.8-31.2, and 3.60-5.20 g/l, respectively. Corresponding milk aggregation properties were in agreement with previous studies [14,18] (data not shown). All samples were centrifuged at 3,000 rpm for 60 min, at 4 °C; resultant fat layers were removed and washed for 4 times with PBS containing a protease inhibitor cocktail (Sigma) [23]. Skimmed milk samples were also added with the same protease inhibitor cocktail before further treatment [22].

To recover whole proteins, a part of skimmed milk samples (20 ml) was directly precipitated with chloroform/methanol [27] and vacuum dried. To isolate whey proteins, the remaining material from each skimmed milk sample (780 ml) was added with 10% v/v acetic acid to reach a final pH value of 4.6; precipitated caseins were then removed by centrifugation at 10,000 rpm for 15 min, at 4 °C. In all cases, resulting whey material was then added with the limited amount of proteins deriving from the corresponding precipitated skimmed milk to generate the fraction containing water-soluble components. In parallel, fat layer samples were dissolved in chloroform/methanol, centrifuged at 12,000 rpm for 5 min, at 4 °C, and removed for their upper phase; then, they were washed with methanol (3 vol) and centrifuged again at 12,000 rpm for 5 min, at 4 °C. Pellets containing whole MFG proteins were finally vacuum dried.

Soluble fraction samples were added with 12 mM dry monopotassium phosphate along with 75 mM NaCl, so as to get the equivalent of PBS, pH 7.2. Then, they were mixed with 0.5 ml of ProteoMiner™ resin (BioRad) and shaken at room temperature, overnight [22]. After removing the supernatant, the resin samples were washed twice with PBS for removal of the excess of soluble proteins. Captured proteins were eluted by three consecutive treatments of the resin samples with 300 ml of 10% w/v SDS, 3% w/v DTE, followed by boiling for 10 min and centrifugation at 4,000 rpm for 2 min [22]. For each sample, protein desorption eluates were pooled and SDS was

eliminated by chloroform/methanol precipitation [27]. Proteins recovered from the enriched soluble and the MFG fractions were quantified with the Bradford colorimetric method [28] and stored at −80 °C until used. The experimental scheme used in this study for the preparation and the analysis of milk proteins is summarized in Fig. 1.

SDS-PAGE and protein digestion

Protein samples from equalized soluble and MFG fractions were added with DTT, boiled and resolved by SDS-PAGE on a 9–16%T acrylamide gradient gel, which was stained with colloidal Coomassie staining [23]. Whole gel lanes were cut into 6 slices, minced and washed with water. Corresponding proteins were *in-gel* reduced, alkylated with iodoacetamide and digested with trypsin. A fourth of each tryptic digest was directly analyzed by nLC-ESI-LIT-MS/MS, while the remaining part was subjected to enrichment for diol-containing compounds by chromatography on a phenylboronic acid-functionalyzed resin [22,23].

Phenylboronate chromatography

m-Aminophenylboronic acid-agarose (Sigma) was suspended in washing buffer (250 mM ammonium acetate, 50 mM $MgCl_2$, pH 8) and packed into a 1 ml polypropylene columns (85 x 7.5 mm) (Qiagen, Germany) [22]. The columns were equilibrated with washing buffer before tryptic digests (dissolved in 200 ml of washing buffer) were loaded at 4 °C. After 1 h, the columns were washed with 10 ml of washing buffer, at 25 °C; modified peptides were finally eluted with 5 ml of 250 mM acetic acid, pH 2.8 [22].

Mass spectrometry analysis of lactosylated and AGEs-containing peptides

Tryptic digests from whole gel portions or eluted fractions from *m*-aminophenylboronic acid-agarose chromatography (enriched for diol-containing peptides) were analyzed by nLC-ESI-LIT-

MS/MS using a LTQ XL mass spectrometer (Thermo, San Jose, CA, USA) equipped with a Proxeon nanospray source connected to an Easy-nanoLC (Proxeon, Odense, Denmark) [22,23]. Peptides were resolved on an Easy C18 column (100 × 0.075 mm, 3 μm) (Proxeon) using a chromatographic gradient of ACN containing 0.1% formic acid in aqueous 0.1% formic acid; gradient was initiated 20 min after sample loading; ACN ramped from 5% to 40% over 90 min, from 40% to 60% over 10 min, and from 60% to 95% over 20 min, at a flow rate of 300 nl/min. Each sample was analyzed on the same LC column under two ETD data acquisition conditions. Acquisition methods were the following: i) data-dependent product ion scanning procedure over the 5 most abundant ions, enabling dynamic exclusion (repeat count 1 and exclusion duration 60 s), and subsequent ETD fragmentation of peptide precursor ions; mass isolation window and ETD activation time were *m/z* 3 and 100 ms (plus supplemental activation), respectively; ii) as the previous method, with exclusion of the peptides confidently identified in the first run, in order to increase the number of the peptides characterized.

Modification site assignment

For identification of lactosylated and AGEs-containing peptides, nLC-ESI-LIT-MS/MS raw data files were searched with MASCOT (Matrix Science, UK) present within the Proteome Discoverer software package version 1.0 SP1 (Thermo) using an updated mammalian non-redundant sequence database (UniProtKB release 2013_08 18/09/2013). Database searching was performed by selecting carbamidomethylation at Cys (+57 Da) as fixed modification, oxidation at Met (+16 Da), plus other glycation/glyco-oxidations at Arg/Lys as variable modifications [2-5,8-10,15,19,22,23,29-38]. The latter modifications included: i) formation of the Amadori compound with lactose (Lact) (+324 Da) or glucose/galactose (Hex) (+162 Da), together with their mono-dehydrated (Lact-H_2O and Hex-H_2O at +306 and +144 Da, respectively) and bis-dehydrated (Lact-$2H_2O$ and Hex-$2H_2O$ at +288 and +126 Da, respectively) derivatives, carboxymethylation (+58 Da),

carboxyethylation (+72 Da), formation of PYR (+108 Da) and 1-alkyl-2-formyl-3,4-glycosyl-pyrrole (AFGP) (+270 Da) at Lys; ii) formation of G-H (+40 Da), MG-H (+54 Da), triosone-derived hydroimidazolone (Trios-H) (+70 Da), 3-DPen-H (+114 Da), 3-DG-H (+144 Da), 3-DGal-H (+144 Da), 3-deoxylactosone-derived hydroimidazolone (3-DLact-H) (+306 Da), G-DH (+58 Da), MG-DH (+72 Da), triosone-derived dihydroxyimidazoline (Trios-DH) (+88 Da), 3-DPen-DH (+132 Da), 3-DG-DH (+162 Da), 3-DGal-DH (+162 Da), 3-deoxylactosone-derived dihydroxyimidazolines (3-DLact-DH) (+324 Da), glucosone-derived dihydroxyimidazolines (Gluc-DH) (+178 Da), galactosone-derived dihydroxyimidazolines (Gal-DH) (+178 Da), DHP (+126 Da), THP (+144 Da) and Arg-pyrimidine (RPYR) (+80 Da) at Arg; iii) formation of G-He (+58 Da), MG-He (+72 Da), triosone-derived hemiaminal (Trios-He) (+88 Da), 3-DPen-He (+132 Da), 3-DG-He (+162 Da), 3-DGal-He (+162 Da), glucosone-derived hemiaminal (Gluc-He) (+178 Da), galactosone-derived hemiaminal (Gal-He) (+178 Da), 3-deoxylactosone-derived hemiaminal (3-DLact-He) (+324 Da) and lactosone-derived hemiaminal (Lact-He) (+340 Da) at Arg or Lys. These linear glycation and glyco-oxidation products are summarized in Supplementary Fig. S2. Database searching also included variable formation of intramolecular cross-linked peptides containing: i) glucose/galactose-associated crossline (+253 Da), lactose-associated crossline (+415 Da), glucose/galactose-associated fluorolink (+235 Da), lactose- associated fluorolink (+397 Da), glyoxal-derived lysine dimer (GOLD) (+35 Da), methylglyoxal-derived lysine dimer (MOLD) (+49 Da), 3-deoxypentosone-derived lysine dimer (POLD) (+109 Da), 3-deoxyglucosone/3-deoxygalactosone-derived lysine dimer (DOLD) (+139 Da), glucosone/galactosone-derived lysine dimer (GLUCOLD) (+155 Da), 3-deoxylactosone-derived lysine dimer (DLactOLD) (+301 Da) and lactose-derived lysine dimer (LACTOLD) (+317 Da) moieties bridging different Lys residues; ii) glucosepane (+108 Da), pentosidine (PENT) (+59 Da), 3-glyoxal-derived imidazolium cross-link (GODIC) (+22 Da), methylglyoxal-derived imidazolium cross-link (MODIC) (+36 Da), 3-deoxypentosone-derived imidazolium cross-link (DPenDIC) (+96 Da), 3-deoxyglucosone/3-

deoxygalactosone-derived imidazolium cross-link (DOGDIC) (+126 Da) and 3-deoxylactosone-derived imidazolium cross-link (DLactDIC) (+288 Da) moieties bridging Arg and Lys residues (Supplementary Fig. S2). Mass searches were carried out by using a mass tolerance value of 2.0 Da for precursor ion and 0.8 Da for MS/MS fragments, trypsin and/or slymotrypsin as proteolytic enzymes, a missed cleavage maximum value of 4. Peptide modification assignment was always associated with manual spectral verification. Data sets were also searched against a reversed protein sequence database, and raw results were filtered under the same criteria; calculated false discovery rate provided a median value of 1.3%. Reported are uniquely those modified peptides common to the three samples analyzed for each type of milk, where the modification was unequivocally assigned to a specific residue. To avoid data redundancy, the modified species showing the highest values of identification scores were reported.

Results and Discussion

Raw, pasteurized, UHT, and powdered for infant nutrition milk samples were skimmed and then directly precipitated to isolate more abundant proteins (1/40 of the whole material) or treated under acid conditions to remove caseins (39/40 of the whole material). For each milk sample, corresponding whey material and precipitated skimmed milk proteins were then mixed together, and the resulting material (soluble fraction) was enriched for less abundant proteins by using a combinatorial peptide ligand library [22]. In parallel, the different milk samples was treated to isolate corresponding MFG proteins [23]. Thus, proteins recovered from the various enriched soluble and MFG fractions were quantified and resolved by SDS-PAGE (Fig. 2). Identical electrophoretic results were observed for the three commercial samples analyzed for each kind of milk investigated (data not shown). Electrophoretic patterns did not allow the recognition of distinct bands associable to cross-linked proteins; this was evident for the MFG fraction that was demonstrated to have a modified protein composition as result of the milk processing conditions

[23]. Whole gel lanes were cut into slices, which were treated with reducing and alkylating reagents, and then with trypsin. To identify lactosylated/AGEs-containing peptides, a part of each tryptic digest was directly analyzed by nLC-ESI-LIT-MS/MS with ETD fragmentation. To enrich for diol-containing modified peptides, the remaining material was subjected to phenylboronate-based chromatography before nLC–ESI–LIT–MS/MS with ETD (Fig. 1).

With the aim to provide a comprehensive analysis of the AGEs-containing proteins present in the different milk products, we used a modified version of the shotgun approach we experienced successful in determining the corresponding lactosylation counterparts [22,23]. To this purpose, a list of potential non-enzymatic glycation and glyco-oxidation modifications (Supplementary Fig. S2) was compiled and used for MS and MS/MS data-driven database searching. Some of these amino acid modifications have already been detected in milk proteins [2-5,13,15,17,19,20]; others have been identified in other model proteins/peptides after their heating in the presence of sugars and/or sugar oxidation products [8-10,15,29-38].

MS analysis of the peptide mixtures was performed by using ETD fragmentation. This was due to the previous experience of us and other authors during MS analysis of lactosylated peptides [1,22,23], and preliminary, comparative ETD and CID fragmentation experiments we performed on the corresponding Amadori adducts with glucose/galactose, their dehydrated counterparts, and G-DH, Trios-DH, CML and PYR derivatives. As an example, Fig. 3 shows the comparison of the CID and ETD spectra recorded for the triply charged ion of the modified peptide 48-64 from glycosylation-dependent cell adhesion molecule 1, bearing a G-DH adduct at Arg56. Under CID fragmentation, sequence-informative b and y ions resulting from peptide backbone cleavage were rare in the spectrum and poorly identified with confidence (Fig. 3A). The most abundant ions were associated with neutral losses of various H_2O molecules; CID analysis of doubly and quadruply charged ions showed that the neutral loss pattern of variable number of water molecules was generally charge-independent. When the triply charged ion was subjected to ETD fragmentation, a

295 series of clear c- and z-type ions was observed (Fig. 3B). A reduced amount of fragment ions
296 corresponding to neutral losses of H_2O molecules from the G-DH adduct was also present,
297 demonstrating that the dihydroxyimidazoline derivative is more stable under ETD fragmentation
298 conditions. Experiments on other G-DH modified peptides demonstrated that the abundance and the
299 almost completeness of the c- and z-type ion series detected, regardless of whether the modification
300 site was located in the middle of the sequence or close to the N-terminus (data not shown), greatly
301 facilitated peptide sequencing and modification site assignment. Similar spectral profile differences
302 were also observed for the corresponding triosone-DH-containing peptide homologue
303 (Supplementary Fig. S3). Dissimilarities in the abundance of sequence-informative fragment ions,
304 as deriving from CID- or ETD-based fragmentation, were also evident for the dehydrated,
305 lactosylated peptides. For example, Supplementary Fig. S4 shows the data recorded for the
306 modified peptide 78-91 from β-LG, bearing a Lact-H_2O adduct at Lys83. In this case, the CID
307 spectrum was populated by fragments deriving from the cleavage of the glycosidic bond, plus
308 neutral losses of molecules of H_2O, and showed rare b- and y-type ions. Conversely, sequence-
309 informative c- and z-type ions resulting from the peptide backbone cleavage were evident in the
310 corresponding ETD spectrum and were identified with confidence.

311 A general similarity in the information provided by both fragmentation methods was observed
312 during the comparative MS analysis of peptides containing CML and PYR adducts. In this context,
313 Supplementary Fig. 5 reports the CID and ETD spectra of the CML-containing peptide 125-138
314 from β-LG, where a comparable number of sequence-informative fragment ions was detected. In
315 both cases, modification was assigned at Lys135. On the basis of what reported above and the
316 consistent enrichment of diol-containing AGEs structures after phenylboronate chromatography
317 [22-24], ETD was then chosen as the preferential fragmentation technique.

318 In general, a number of AGEs-containing peptides were attributed unambiguously to the
319 modification of specific Lys or Arg residues, based on the sequence, the unique mass shift and

fragment ions. In this context, the use of trypsin for enzymatic digestion resulted efficient in limiting the number of peptides bearing simultaneously Arg and Lys within their sequence. In other cases, AGEs assignment to specific residues was not definitive, since the same mass shift value can be associated with different adducts at the same amino acid. Supplementary Information reports the cases in which this ambiguity occurred. Nevertheless, it has to be mentioned that the Amadori compounds and CML, CEL and DH derivatives are generally more stable the various isobaric hemiaminal counterparts [10,19].

Database searching for lactosylated peptides from mammalian proteins revealed a perfect agreement with the data already published in our previous studies on commercial milk products [22,23], demonstrating a high reproducibility of the shotgun approach we used. In fact, 271 non-redundant lactosylation sites were confirmed in 33 proteins from the soluble milk fractions (data not shown). Similarly, 157 modification sites affected by lactosylation were verified in 35 MFG proteins, in addition to the 153 ones present in other 21 soluble proteins detected as adsorbed on globule material (data not shown). Notably, the present study originally identified: i) 12 non-redundant lactosylation sites in 10 proteins never observed as lactosylated in milk; ii) 28 non-redundant lactosylation sites in 20 proteins already observed as lactosylated in milk, but at different Lys residues; iii) 36 non-redundant lactosylation sites that were not detected in a specific thermally-treated milk, but were already detected in other commercial products [22,23]; iv) 15 non-redundant lactosylation sites that were detected in a specific milk fraction (whey or MFGs), contrarily to what detected before (MFGs or whey) [22,23]. Supplementary Table S1 summarizes the novel lactosylation sites described in this study, together with the modified peptides that were already described as lactosylated in our previous investigations [22,23] and were detected here as bearing various AGE adducts (see below). As expected, a progressive increase in the number of lactosylation sites was observed in milk proteins, which well paralleled the harshness of food processing conditions [17,19,22,23].

More importantly, 289 modified peptides bearing: i) Lact-H_2O (+306 Da), Lact-$2H_2O$ (+288 Da), the Amadori compound with glucose or galactose/3-DG-He/3-DGal-He (+162 Da), Hex-H_2O (+144 Da), Hex-$2H_2O$ (+126 Da), CML/G-He (+58 Da), CEL/MG-He (+72 Da), PYR (+108 Da), 3-DPen-He (+132 Da) and Lact-He (+340 Da) adducts at Lys residues; ii) RPYR (+80 Da), G-H (+40 Da), MG-H (+54 Da), G-DH/G-He (+58 Da), MG-DH/MG-He (+72 Da), 3-DG-DH/3-DGal-DH/3-DG-He/3-DGal-He (+162 Da), 3-DLact-DH/3-DLact-He (+324 Da), 3-DLact-H (+306 Da) and Trios-DH/Trios-He (+88 Da) adducts at Arg, were recognized in the different milk samples. As mentioned above, AGEs assignment in some cases was not definitive, since the same mass shift value can be associated with different products occurring at a specific residue (Supplementary Fig. 2). On the other hand, recent studies on hemiaminals have reported these compounds as being not very stable for prolonged times [10,19]; accordingly, they can be eventually excluded when other stable compounds may occur together. Globally, 3, 30, 112 and 281 non-lactosylated modified species were detected in raw, pasteurized, UHT, and powdered for infant nutrition milk samples, respectively. They were associated with 169 non-redundant intermediate- and advanced glycation end-products-modified sites in 30 well known milk proteins that (with a unique exception) have been already observed as lactosylated in this fluid, strengthening the consistency of our data. Their nature and corresponding MS identification details are reported in Supplementary Tables S1-S9. Most frequent intermediate and advanced glycation end-products were the Amadori compounds with glucose or galactose/3-DG-He/3-DGal-He, Lact-H_2O, CML/G-He, G-DH/G-He, 3-DG-DH/3-DGal-DH/3-DG-He/3-DGal-He and 3-DLact-DH/3-DLact-He derivatives (about 52%, 10%, 8%, 5%, 4% and 3% of the whole modified, non-lactosylated species reported here, respectively). Additional peptide adducts (16 in number) showing a Δm = +40 Da or +54 Da assigned to Lys residues were also observed. These derivatives were tentatively associated with the Schiff bases of GO and MGO (G-SB and MG-SB, respectively) [39], which modified 17 sites in a total of 9 proteins. The distribution of the non-lactosylated, modified peptides detected in the different

370 fractions of the various milk products here assayed is reported in Fig. 4; only those components
371 common to all samples of each milk type are reported in Supplementary Tables S1-S9).

372 Notably, detection of intermediate and advanced glycation end-products-containing peptides
373 in the different milk fractions well paralleled with the identification of the corresponding protein
374 species therein [22,23] and the simultaneous recognition of the corresponding lactosylated peptides
375 in the same sample or in milk samples subjected to softer processing conditions (Supplementary
376 Table S1). These findings confirmed that, according to the individual reactivity of lysines present in
377 proteins and to the harshness of thermal treatment exerted, modifiable Lys residues firstly react with
378 the most abundant sugar present in milk, i.e. lactose, to generate the corresponding Amadori
379 compounds. Then, these early modification products are further modified/oxidized to generate the
380 corresponding intermediate and advanced glycation counterparts. On the other hand, non-reacted
381 Arg and Lys residues in these proteins can be modified by minor sugars present in milk, i.e. glucose
382 or galactose, or by lactose degradation/oxidation products generated following thermal treatment,
383 i.e. G, MG, 3-DLact, 3-DG, 3-DGal, 3-DPen, lactosone and triosone, ultimately leading to the
384 production of the corresponding AGEs. Progressively augmented detection of intermediate and
385 advanced glycation end-products in milk samples subjected to harshening food processing
386 conditions was in good agreement with quantitative data from GC-MS or LC-MS analysis of amino
387 acid adducts, as reported for commercial milk samples or related model systems [2,5,40].

388 A comparison of the data here reported for β-LG, α-lactalbumin (α-LA), and various casein
389 (CN) species with that recently published in other investigations on specific proteins from raw and
390 commercial milk samples and model proteins subjected to thermal treatment in the presence of
391 sugars demonstrated their good concurrence either in terms of the nature of the AGEs observed in
392 the different products and of the modified Lys/Arg residues present therein [4,13,15-17,19,35].
393 Eventual discrepancies between these studies may be ascribed to the phenylboronate

chromatography we used in this work, which promoted a general enrichment/overestimation of diol-containing AGEs-modified peptides in all milk fractions.

Our study also revealed the occurrence of three linear peptides bearing an intra-molecular cross-linking between Lys and Lys/Arg residues, thus generating pentosidine, DOGDIC and DOLD structures (Supplementary Table S1). As an example, Fig. 5 shows the ETD spectrum recorded for the cross-linked peptide 153-165 from α-S2 casein, which presents a pentosidine adduct linking together Lys158 and Arg160. The concomitant assignment of fragment ions to both bridged amino acids ensured no ambiguity in the identification of this cross-linked peptide adduct. Similarly, a DOLD adduct bridging Lys114 and Lys115 was observed for the cross-linked peptide 113-123 from serotransferrin (Supplementary Fig. S6). Also in this case, the ETD spectrum did not provide uncertainty in peptide assignment, due to the parallel recognition of fragment ions to both bridged Lys residues. Some of the intra-molecular cross-linking products reported above have been already described in dedicated studies on model proteins heated in the presence of sugars [32,37,38] or on biomarker components of specific pathologies [34,36]. The nature of the observed cross-linked species and the frequency of the other linear AGEs detected in this study suggest G, 3-DG and 3-DGal as the main reactive dicarbonyl compounds in thermally-processed milk samples. This observation is in perfect agreement with quantitative data from a dedicated investigation on the reactive species in milk products that result from the oxidative degradation of lactose [7].

A visual inspection of the spectroscopic/crystallographic structure of the milk proteins reported in Supplementary Table S1 revealed that most of the labeled residues occurs on their macromolecule surface (data not shown), although few internal amino acids were also affected. Accordingly, the Maillard reaction can influence the 3D structure of these proteins, as already observed in various model systems [41], with significant consequences on their functionality. Since it has been shown that various milk proteins are resistant to proteolysis in the gastrointestinal tract [42,43], and may directly exert (even in low copies) a number of biological activities within the

infant intestine or other organs [43-45], their widespread glycation/glyco-oxidation may have important consequences on food nutritional and health-beneficial characteristics. Here, we have demonstrated that various milk proteins involved in the delivery of nutrients, the defence response against pathogens, related inflammatory processes, and the regulation of cellular proliferative/differentiation events [22,23,44,45] are subjected to glycation/glyco-oxidation after thermal processing (Supplementary Table S1). Thus, a severe heat treatment can influence their activity [46,47]. This matter can have important consequences when milk is the unique source of nutrients/protective molecules, as in infant diet.

Since modified Lys/Arg residues are no longer recognized by proteolytic enzymes, the Maillard reaction can also affect the gastrointestinal digestion of milk proteins [48]. This issue was recently addressed by comparing the digestibility of the milk proteins present in different commercial samples [14,23]. When major soluble proteins were evaluated, an increased *in vitro*/*in vivo* digestion resistance of the intact species was observed in raw and pasteurized samples by SDS-PAGE, compared to the UHT and sterilized ones. However, protein digestibility (as determined by a Kjedahl procedure) showed a slight opposite trend, as already reported for milk-based infant formulas [49,50]. This effect was interpreted assuming that industrial heating can improve the digestibility by protein denaturation, but this enhancement is likely to be offset by heat-derived modifications involved in decreasing proteolysis of milk components [14]. When MFG proteins were evaluated, an increasing resistance to trypsin proteolysis was observed moving from pasteurized to UHT and powdered milk proteins [23]. In the whole, these results confirmed previous studies on model milk proteins treated with various sugars/oligosaccharides [51,52]. Not considering possible physiological implications on the eventual assumption of poorly-digestible foods, this point has a first important significance for some soluble milk proteins that, after an impaired proteolysis, may present a reduced release of well-known sequence-encrypted bioactive peptides [53,54]. A reduced biological activity of these components can also be eventually ascribed

444 to non-desired peptide glycation/glyco-oxidation. On the other hand, a reduced digestibility of
445 modified milk proteins may affect the allergic response to specific dairy products; in fact, it has
446 been reported that proteins resistant to digestion may behave as major food allergens [55,56].
447 Extensive glycation/glyco-oxidation of surface-exposed residues in specific milk proteins should
448 also *per se* modulate the allergenic impact of these nutrients [26,57,58].
449
450 **Conclusions**
451 Although fundamental progresses have been made in research in the field of the Maillard
452 reaction, basic questions from a chemical, biological and physiological point of view are still
453 unanswered. Without doubt, we can say that the reaction products we know today do only represent
454 a part of the total AGEs probably present in ‘‘browned’’ foods. Following a preliminary phase
455 realized on model protein systems, where the main Maillard reaction products were characterized,
456 novel efforts nowadays have to be spent in identifying and quantifying individual protein
457 derivatives in food matrices, through reliable analytical techniques. In this context, MS-based
458 approaches have been recently used to fully characterize the glycating effectors of important AGEs-
459 precursors and their products in model peptides [10,59], the glycation/glyco-oxidation products in
460 specific proteins from complex foods [16-18] or the main modification protein targets (including
461 modification sites) therein, according to a proteomic perspective [13-15,19,22,23,60]. Due to the
462 accumulated experience in this research field, a large part of these studies have been performed on
463 milk products. Focusing on milk, we believe that the present study will add novel information to the
464 intricate scenario of the compounds deriving from the Maillard reaction on milk proteins and
465 carbohydrates, by providing the widest inventory of assigned lactosylation and AGEs-containing
466 species in commercial products determined so far. A number of AGEs previously undetected in
467 milk proteins is here also originally described. This information is important for its nutritional
468 consequences since AGEs-containing milk proteins are daily ingested in large amounts from

humans and, as unique polypeptide source, from infants. Although lacking quantitative information, these data will help in making comprehensive milk-focused database, which may ultimately allow food manufacturers to thermally process their products as careful as possible in order to obtain products with controlled amounts of AGEs. To reach this ultimate goal, future studies applying quantitative MS procedures on isolated protein components modified *in vitro* with unlabelled/isotopically-labeled lactose mixtures or quantitative proteomic approaches (based on labeled proteotypic peptides or AGEs-oriented derivatizing reagents) will be advisable to gain quantifiable data on the amino acids more prone to generate adducts in various milk proteins [1,9,20,61]. These information are important points of interest to connect the extent of the Maillard reaction occurring in different commercial samples with potential nutritional implications. The themes described above and those associated with the other modification reactions present as a consequence of milk industrial treatments, which determine the generation of additional oxidation and advanced lipooxidation end-products [2,12,61], have to be fully evaluated for a whole estimation of the nutritional/toxicological properties of dairy products deriving from severe heat processing.

Acknowledgements

This study was supported by grants from the Italian Ministry of Economy and Finance for the project "Innovazione e sviluppo del Mezzogiorno - Conoscenze Integrate per Sostenibilità ed Innovazione del *Made in Italy* Agroalimentare - Legge n. 191/2009" and from the Campania Region for the project "QUARC: qualità delle produzioni tipiche campane ed il loro territori – Progetto di ricerca industriale e sviluppo sperimentale per la realizzazione di Campus dell'Innovazione - O.O. 2.1/2.2 P.O. FESR Campania 2007/2013".

References

1 Siciliano RA, Mazzeo MF, Arena S, Renzone G, Scaloni A. Mass spectrometry for the analysis of protein lactosylation in milk products. *Food Res Int* 2013, *54*, 988-1000.

2 Pischetsrieder M, Henle T. Glycation products in infant formulas: chemical, analytical and physiological aspects. Amino Acids 2012,42:1111-18.

3 Henle T. Protein-bound advanced glycation endproducts (AGEs) as bioactive amino acid derivatives in foods. Amino Acids 2005,29:313-22.

4 Meltretter J, Seeber S, Humeny A, Becker CM, Pischetsrieder M. Site-specific formation of Maillard, oxidation, and condensation products from whey proteins during reaction with lactose. J Agric Food Chem 2007,55:6096-103.

5 Ahmed N, Mirshekar-Syahkal B, Kennish L, Karachalias N, Babaei-Jadidi R, Thornalley PJ. Assay of advanced glycation endproducts in selected beverages and food by liquid chromatography with tandem mass spectrometric detection. Mol Nutr Food Res 2005,49:691-9.

6 Szmuda M, Glomb MA. Fragmentation pathways during Maillard-induced carbohydrate degradation. J Agrig Food Chem 2013,61:10198-208.

7 Hellwig M, Degen J, Henle T. 3-deoxygalactosone, a "new" 1,2-dicarbonyl compound in milk products. J Agric Food Chem 2010,58:10752-60.

8 Rabbani N, Thornalley PJ. Glycation research in amino acids: A place to call home. Amino Acids 2012,42:1087-96.

9 Arena S, Salzano AM, Renzone G, D'Ambrosio C, Scaloni A. Non-enzymatic glycation and glycoxidation protein products in foods and diseases: an interconnected, complex scenario fully open to innovative proteomic studies. Mass Spectrom Rev 2014,33:49-77.

10 Mittelmaier S, Pischetsrieder M. Multistep ultrahigh performance liquid chromatography/tandem mass spectrometry analysis for untargeted quantification of glycating activity and identification of most relevant glycation products. Anal Chem 2011,83:9660-68.

11 Mossine VV, Linetsky M, Glinsky GV, Ortwerth BJ, Feather MS. Superoxide free radical generation by Amadori compounds: the role of acyclic forms and metal ions. Chem Res Toxicol 1999,12:230-6.

521 12 Fenaille F, Parisod V, Tabet JC, Guy PA. Carbonylation of milk powder proteins as a
522 consequence of processing conditions. Proteomics 2005,5:3097-104.

523 13 Calvano CD, Monopoli A, Loizzo P, Faccia M, Zambonin C. Proteomic approach based on
524 MALDI-TOF MS to detect powdered milk in fresh cow's milk. J Agric Food Chem
525 2013,61:1609-17.

526 14 Wada Y, Lonnerdal B. Effects of industrial heating processes of milk on site-specific protein
527 modifications and their relationship to *in vitro* and *in vivo* digestibility. J Agric Food Chem
528 2014,62:4175-85.

529 15 Holland JW, Gupta R, Deeth HC, Alewood PF. Proteomic analysis of temperature-dependent
530 changes in stored UHT milk. J Agric Food Chem 2011,59:1837-46.

531 16 Carulli S, Calvano CD, Palmisano F, Pischetsrieder M. MALDI-TOF MS characterization of
532 glycation products of whey proteins in a glucose/galactose model system and lactose-free milk.
533 J Agric Food Chem 2011,59:1793-803.

534 17 Meltretter J, Becker CM, Pischetsrieder M. Identification and site-specific relative
535 quantification of beta-lactoglobulin modifications in heated milk and dairy products. J Agric
536 Food Chem 2008,56:5165-71.

537 18 Scaloni A, Perillo V, Franco P, Fedele E, Froio R, Ferrara L, Bergamo P. Characterization of
538 heat-induced lactosylation products in caseins by immunoenzymatic and mass spectrometric
539 methodologies. Biochim Biophys Acta 2002,1598:30-9.

540 19 Meltretter J, Wüst J, Pischetsrieder M. Comprehensive analysis of nonenzymatic post-
541 translational β-lactoglobulin modifications in processed milk by ultrahigh-performance liquid
542 chromatography-tandem mass spectrometry. J Agric Food Chem 2013,61:6971-81.

543 20 Meltretter J, Wüst J, Pischetsrieder M. Modified Peptides as Indicators for Thermal and Non-
544 Thermal Reactions in Processed Milk. J Agric Food Chem. 2014, in press

545 21 Lima M, Moloney C, Ames JM. Ultra performance liquid chromatography-mass spectrometric
546 determination of the site specificity of modification of beta-casein by glucose and
547 methylglyoxal. Amino Acids 2009,36:475-81.

548 22 Arena S, Renzone G, Novi G, Paffetti A, Bernardini G, Santucci A, Scaloni A. Modern
549 proteomic methodologies for the characterization of lactosylation protein targets in milk.
550 Proteomics 2010,10:3414-34.

551 23 Arena S, Renzone G, Novi G, Scaloni A. Redox proteomics of fat globules unveils broad
552 protein lactosylation and compositional changes in milk samples subjected to various
553 technological procedures. J Proteomics 2011,74:2453-75.

554 24 Zhang Q, Frolov A, Tang N, Hoffmann R, van de Goor T, Metz TO, Smith RD. Application of
555 electron transfer dissociation mass spectrometry in analyses of non-enzymatically glycated
556 peptides. Rapid Comm Mass Spectrom 2007,21:661-6.

557 25 Muscat S, Pischetsrieder M, Maczurek A, Rothemund S, Münch G. Cytotoxicity of Maillard
558 reaction products determined with a peptide spot library. Mol Nutr Food Res 2009,53:1019-29.

559 26 Heilmann M, Wellner A, Gadermaier G, Ilchmann A, Briza P, Krause M, Nagai R, Burgdorf S,
560 Scheurer S, Vieths S, Henle T, Toda M. Ovalbumin modified with pyrraline, a Maillard
561 reaction product, shows enhanced T-cell immunogenicity. J Biol Chem 2014,289:7919-28.

562 27 Wessel D, Flugge UI. A method for the quantitative recovery of protein in dilute solution in the
563 presence of detergents and lipids. Anal Biochem 1984,138:141–3.

564 28 Bradford MM. A rapid and sensitive method for the quantitation of microgramquantities of
565 protein utilizing the principle of protein-dye binding. Anal Biochem 1976,72:248–54.

566 29 Humeny A, Kislinger T, Becker CM, Pischetsrieder M. Qualitative determination of specific
567 protein glycation products by matrix-assisted laser desorption/ionization mass spectrometry
568 Peptide mapping. J Agric Food Chem 2002,50:2153-60.

569 30 Barnaby OS, Cerny RL, Clarke W, Hage DS. Comparison of modification sites formed on
570 human serum albumin at various stages of glycation. Clin Chim Acta 2011,412:277-85.

571 31 Cotham WE, Metz TO, Ferguson PL, Brock JW, Hinton DJ, Thorpe SR, Baynes JW, Ames
572 JM. Proteomic analysis of arginine adducts on glyoxal-modified ribonuclease. Mol Cell
573 Proteomics 2004,3:1145-53.

574 32 Dai Z, Wang B, Sun G, Fan X, Anderson VE, Monnier VM. Identification of glucose-derived
575 cross-linking sites in ribonuclease A. J Proteome Res 2008,7:2756-68.

576 33 Brock JW, Cotham WE, Thorpe SR, Baynes JW, Ames JM. Detection and identification of
577 arginine modifications on methylglyoxal-modified ribonuclease by mass spectrometric
578 analysis. J Mass Spectrom 2007,42:89-100.

579 34 Shao CH, Capek HL, Patel KP, Wang M, Tang K, DeSouza C, Nagai R, Mayhan W, Periasamy
580 M, Bidasee KR. Carbonylation contributes to SERCA2a activity loss and diastolic dysfunction
581 in a rat model of type 1 diabetes. Diabetes 2011,60:947-59.

582 35 Glomb MA, Rösch D, Nagaraj RH. N^{δ}-(5-hydroxy-4,6-dimethylpyrimidine-2-yl)-l-ornithine, a
583 novel methylglyoxal-arginine modification in beer. J Agric Food Chem 2001,49:366-72.

584 36 Sell DR, Biemel KM, Reihl O, Lederer MO, Strauch CM, Monnier VM. Glucosepane is a
585 major protein cross-link of the senescent human extracellular matrix. Relationship with
586 diabetes. J Biol Chem 2005,280:12310-5.

587 37 Biemel KM, Bühler HP, Reihl O, Lederer MO. Identification and quantitative evaluation of the
588 lysine-arginine crosslinks GODIC, MODIC, DODIC, and glucosepan in foods. Nahrung
589 2001,45:210-4.

590 38 Biemel KM, Reihl O, Conrad J, Lederer MO. Formation pathways for lysine-arginine cross-
591 links derived from hexoses and pentoses by Maillard processes: unraveling the structure of a
592 pentosidine precursor. J Biol Chem 2001,276:23405-12.

593 39 Nasiri R, Field MJ, Zahedi M, Moosavi-Movahedi AA. Comparative DFT study to determine if
594 α-oxoaldehydes are precursors for pentosidine formation. J Phys Chem A 2012,116:2986-96.

595 40 Erbersdobler HF, Somoza V. Forty years of furosine - forty years of using Maillard reaction
596 products as indicators of the nutritional quality of foods. Mol Nutr Food Res 2007,51:423-30.

597 41 Boye JI, Alli I. Thermal denaturation of mixtures of β-lactalbumin and α-lactoglobulin: a
598 differential scanning calorimetric study. Food Res Int 2000,33:673-82.

599 42 Castell JV, Friedrich G, Kuhn CSm Poppe GE. Intestinal absorption of undegraded proteins in
600 men: presence of bromelain in plasma after oral intake. Am J Physiol 1997,273:G139-46.

601 43 Lönnerdal B. Nutritional and physiologic significance of human milk proteins. Am J Clin Nutr
602 2003,77:1537S-1543S.

603 44 D'Alessandro A, Scaloni A, Zolla L. Human milk proteins: an interactomics and updated
604 functional overview. J Proteome Res 2010,9:3339-73.

605 45 D'Alessandro A, Zolla L, Scaloni A. The bovine milk proteome: cherishing, nourishing and
606 fostering molecular complexity. An interactomics and functional overview. Mol Biosyst
607 2011,7:579-97.

608 46 Sebekova K, Somoza V. Dietary advanced glycation endproducts (AGEs) and their health
609 effects. Mol Nutr Food Res 2007,51:1079-84.

610 47 Henle T. Dietary advanced glycation end products-a risk to human health? A call for an
611 interdisciplinary debate. Mol Nutr Food Res 2007,51:1075-8.

612 48 Dalsgaard TK, Nielsen JH, Larsen LB. Proteolysis of milk proteins lactosylated in model
613 systems. Mol Nutr Food Res 2007,51:404-14.

614 49 Rudloff S, Lönnerdal B. Solubility and digestibility of milk proteins in infant formulas exposed
615 to different heat treatments. J Pediatr Gastroenterol Nutr 1992,15:25-33.

616 50 Rutherfurd SM, Moughan PJ. Digestible reactive lysine in selected milk-based products. J
617 Dairy Sci 2005,88:40-8.

618 51 Sanz ML, Corzo-Martínez M, Rastall RA, Olano A, Moreno FJ. Characterization and in vitro
619 digestibility of bovine beta-lactoglobulin glycated with galactooligosaccharides. J Agric Food
620 Chem 2007,55:7916-25.

621 52 Corzo-Martínez M, Soria AC, Belloque J, Villamiel M, Moreno FJ. Effect of glycation on the
622 gastrointestinal digestibility and immunoreactivity of bovine beta-lactoglobulin. Int Dairy J
623 2010,20:742-52.

624 53 Clare DA, Swaisgood HE. Bioactive milk peptides: a prospectus. J Dairy Sci 2000,83:1187-95.

625 54 Meisel H. Biochemical properties of peptides encrypted in bovine milk proteins. Curr Med
626 Chem 2005,12:1915-9.

627 55 Astwood JD, Leach JN, Fuchs RL. Stability of food allergens to digestion in vitro. Nat
628 Biotechnol 1996,14:1269-73.

629 56 Wal JM. Structure and function of milk allergens. Allergy 2001,56:Suppl.67,35-8.

630 57 Karamanova L, Fukal L, Kodicek M, Rauch P, Mills E, Morgan M. Immunoprobes for
631 thermally-induced alteration in whey protein structure and their application to the analysis of
632 thermally-treated milks. Food Agric Immunol 2003,15:77-91.

633 58 Gruber P, Vieths S, Wangorsch A, Nerkamp J, Hofmann T. Maillard reaction and enzymatic
634 browning affect the allergenicity of Pru av 1, the major allergen from cherry (*Prunus avium*). J
635 Agric Food Chem 2004,52:4002-7.

636 59 Smuda M, Glomb MA. Novel insights into the Maillard catalyzed degradation of maltose. J
637 Agric Food Chem 2011,59:13254-64.

638 60 Meyer B, Al-Diab D, Vollmer G, Pischetsrieder M. Mapping the glycoxidation product Nε-
639 carboxymethyllysine in the milk proteome. Proteomics 2011,11:420-8.

640 61 Bachi A, Dalle-Donne I, Scaloni A. Redox proteomics: chemical principles, methodological
641 approaches and biological/biomedical promises. Chem Rev 2013,113:596-698.

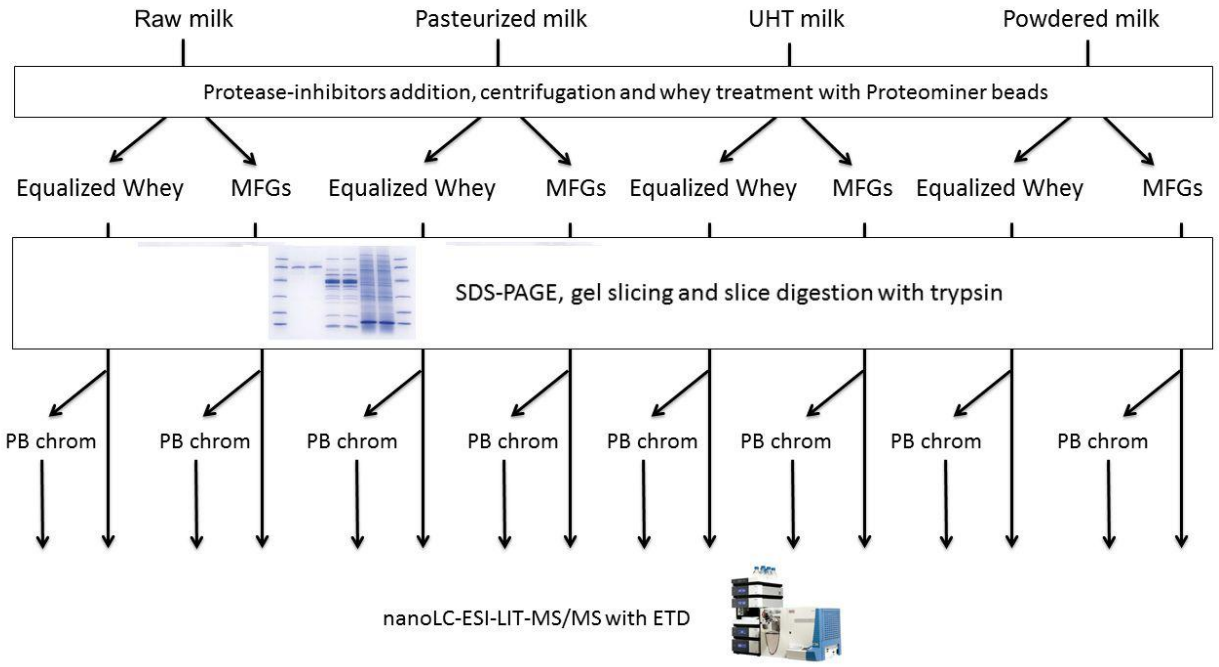
642

643

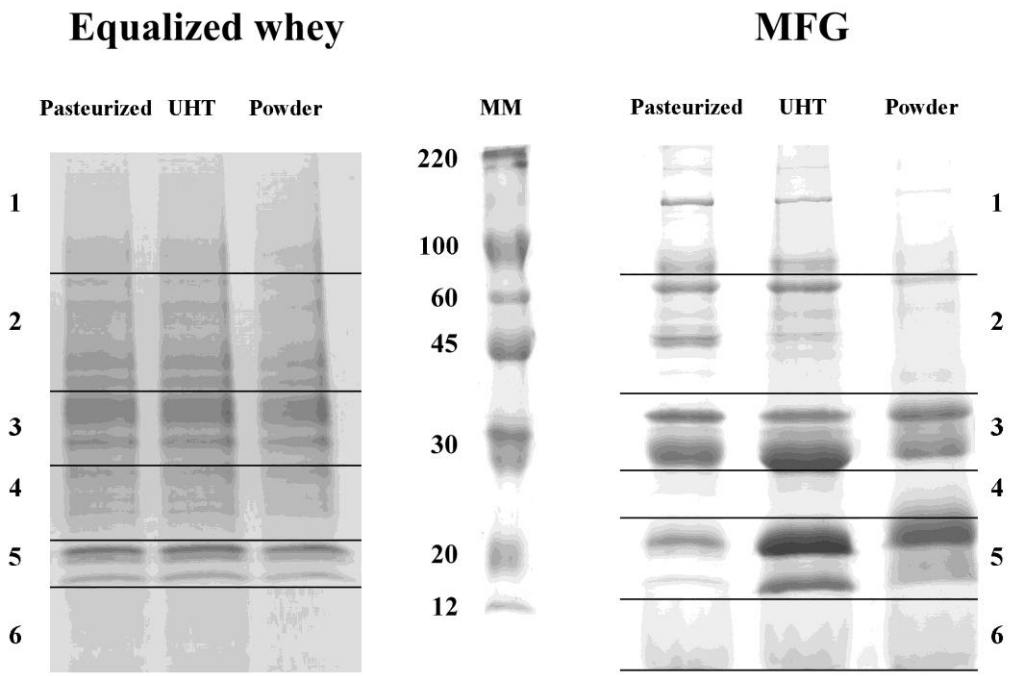
644 **Legend to Figures**

645 **Figure 1.** Schematic diagram illustrating the experimental approach used for the preparation and the
646 analysis of milk proteins. PB chrom, phenylboronate chromatography. Three samples were analyzed
647 for each kind of milk. Whey proteins added with amounts of counterparts from corresponding
648 skimmed milk was used to prepare soluble components that were further subjected to Proteominer-
649 based enrichment; this mixed material is named in the figure as equalized whey.
650
651 **Figure 2.** SDS-PAGE of soluble and fat globule proteins from pasteurized, UHT and powdered milk
652 for infant nutrition. For simplicity, data for a single commercial product of each milk kind are shown.
653 Identical results were observed for the remaining two products assayed (data not shown). Molecular
654 markers (MM) are shown, together with the gel portions (six in number) from each lane that were
655 subjected to proteomic analysis for the characterization of lactosylated and AGEs-modified peptides.
656 Equalized material resulting from starting whey proteins added with amounts of counterparts from
657 corresponding skimmed milk is named in the figure as equalized whey.
658
659 **Figure 3.** CID (A) and ETD (B) MS/MS spectra of the $[M+3H]^{3+}$ ion at *m/z* 685.3 associated with
660 the G-DH-modified peptide (48-64) from glycosylation-dependent cell adhesion molecule 1, where r
661 represents the glyoxal-deriving dihydroxyimidazoline adduct to Arg56. The spectra were acquired
662 with alternating CID and ETD scanning.
663
664 **Figure 4.** Venn diagram showing the distribution of the non-lactosylated, modified peptides as
665 detected in the various raw and commercial samples analyzed in this study. Independent data for the
666 corresponding soluble and MFG fractions are provided. Only modified components common to all
667 three samples analyzed for each kind of milk are reported in Supplementary Tables S1-S9 and in the
668 main text.
669

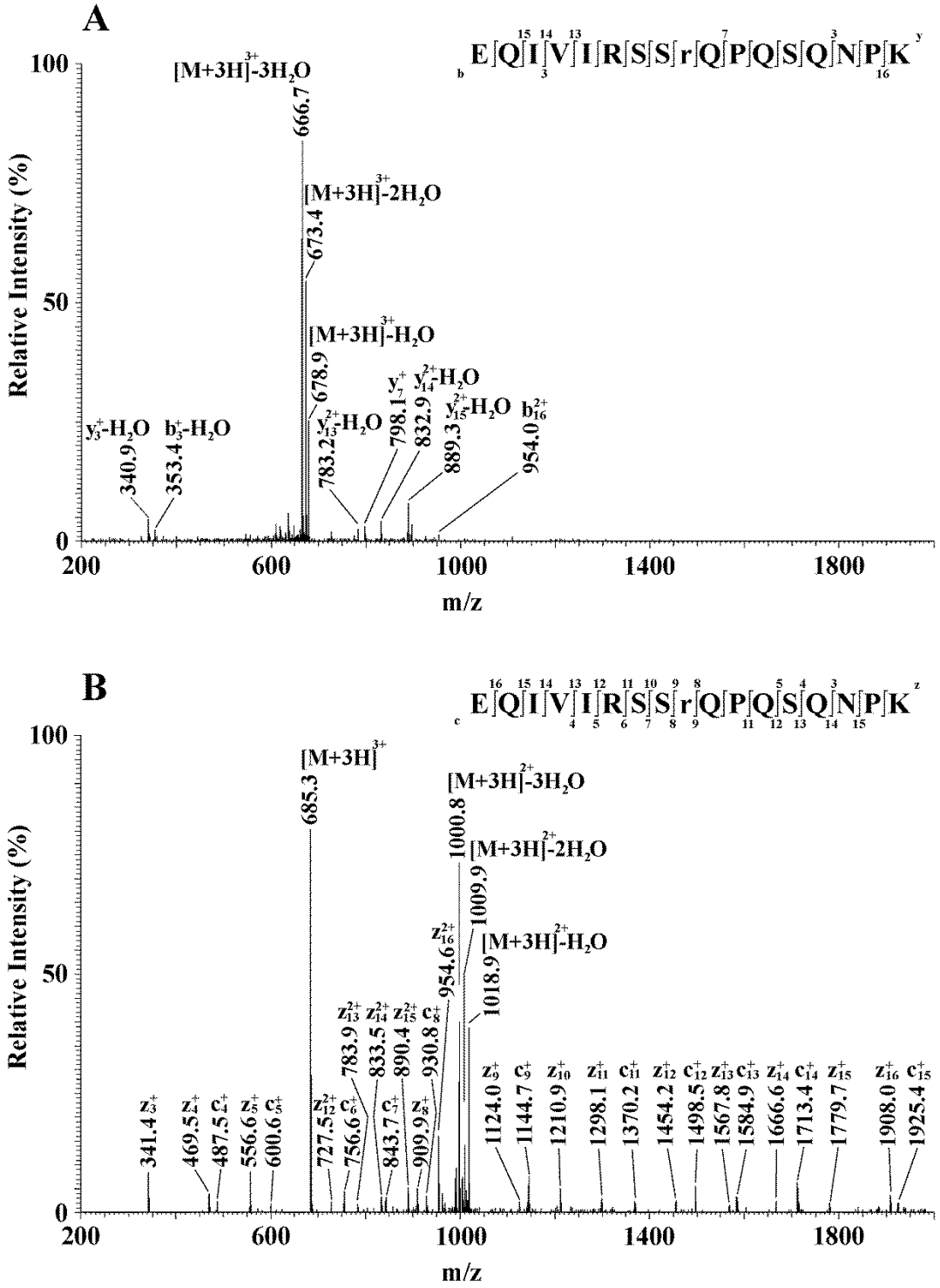
Figure 5. ETD MS/MS spectrum of the $[M+3H]^{3+}$ ion at *m/z* 564.9 associated with the pentosidine-containing peptide (153-165) from alpha-S2 casein, where the pentosidine cross-linking between Lys158 and Arg160 is indicated.



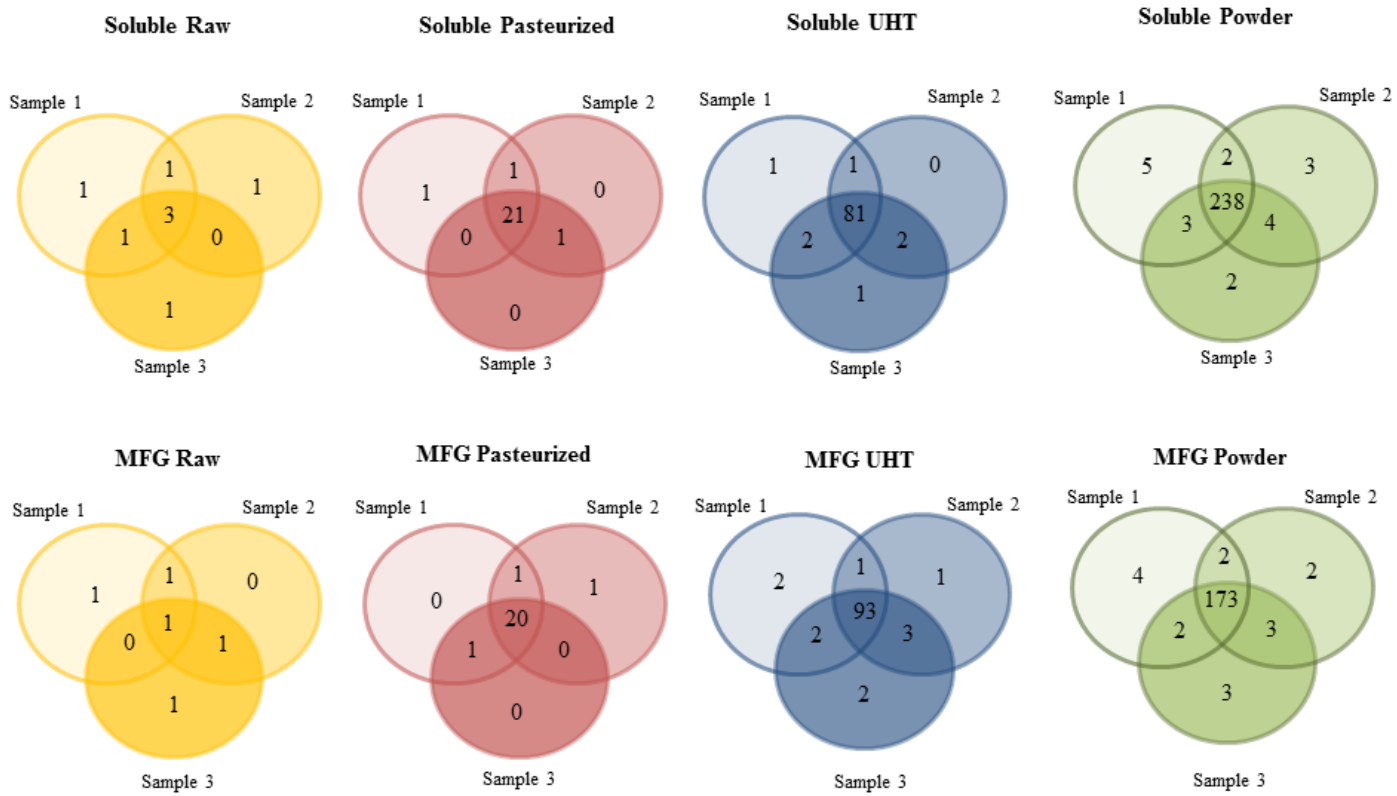
Renzone et al., Figure 1



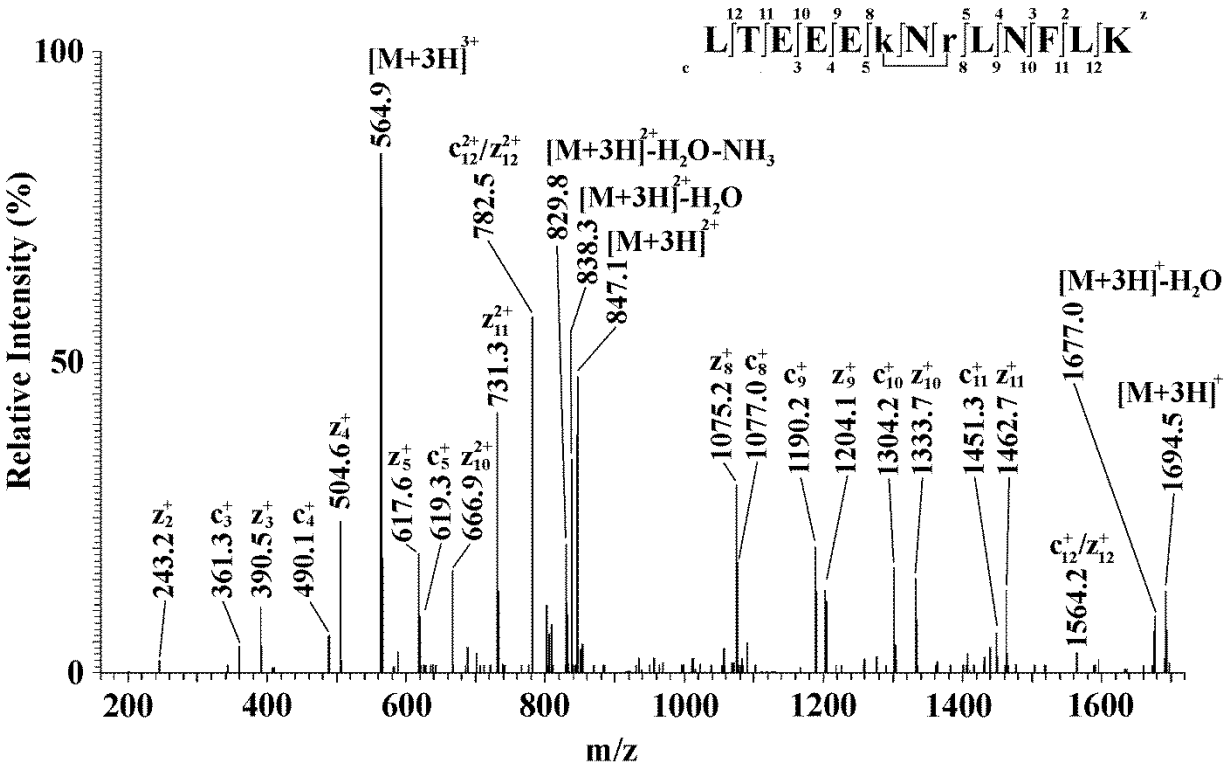
Renzone et al., Figure 2



Renzone et al., Figure 3



Renzone et al., Figure 4



Renzone et al., Figure 5

Figure

Click here to download high resolution image

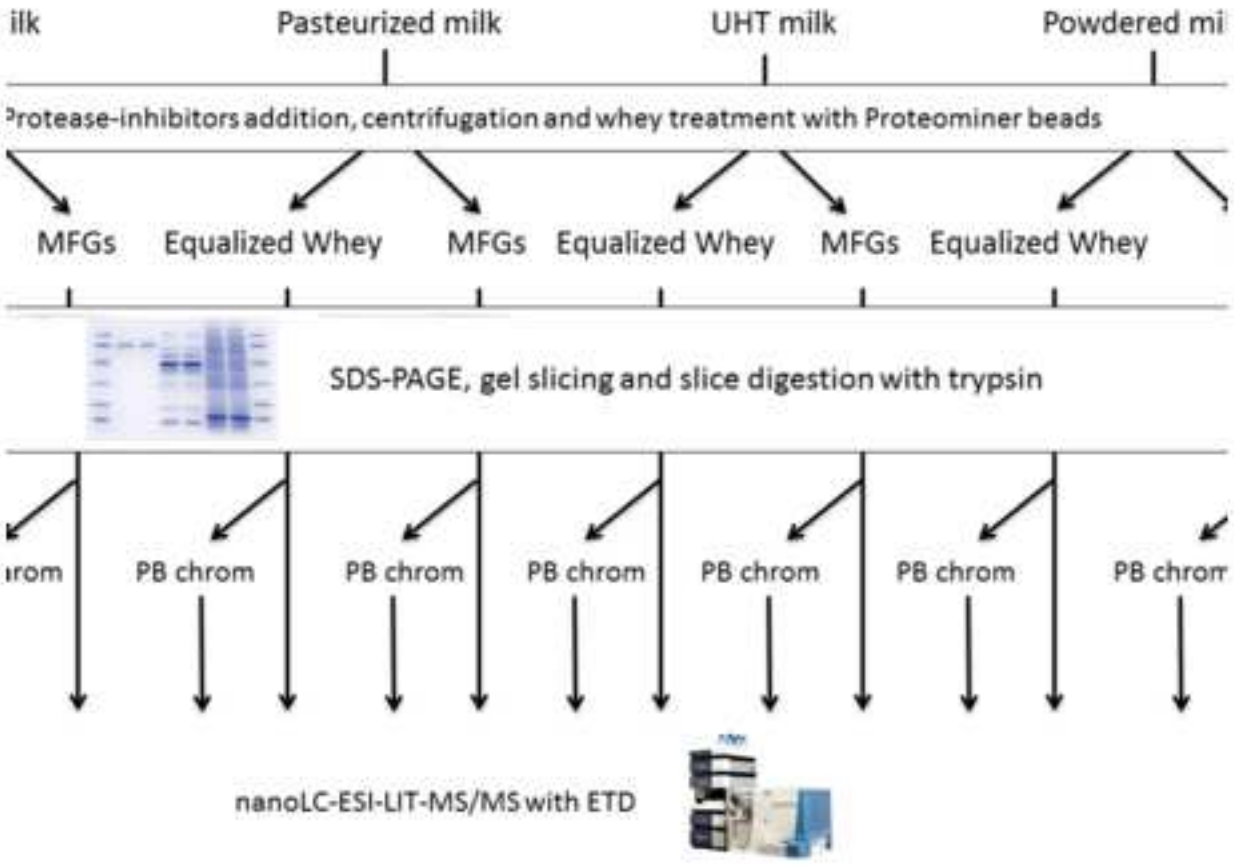
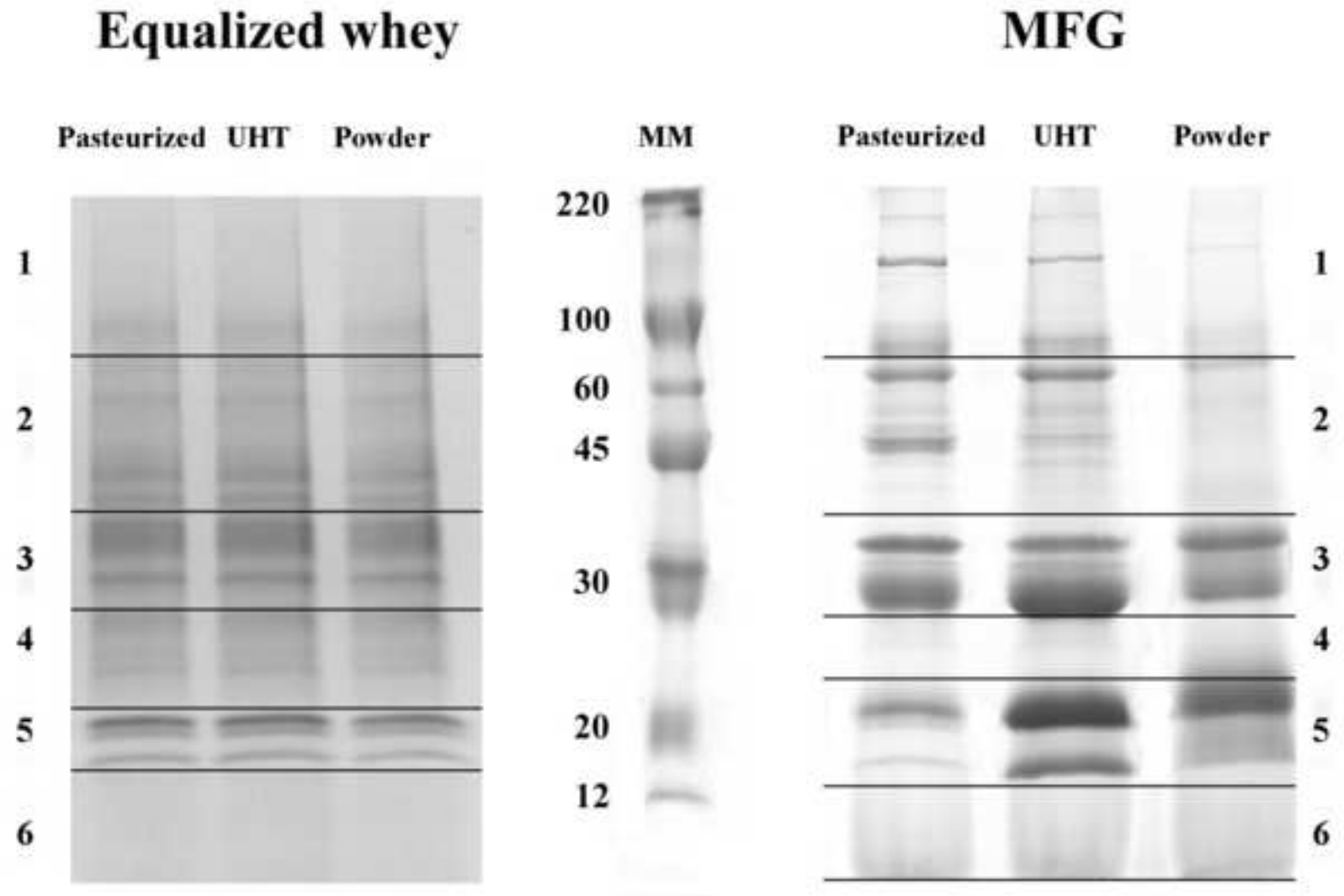
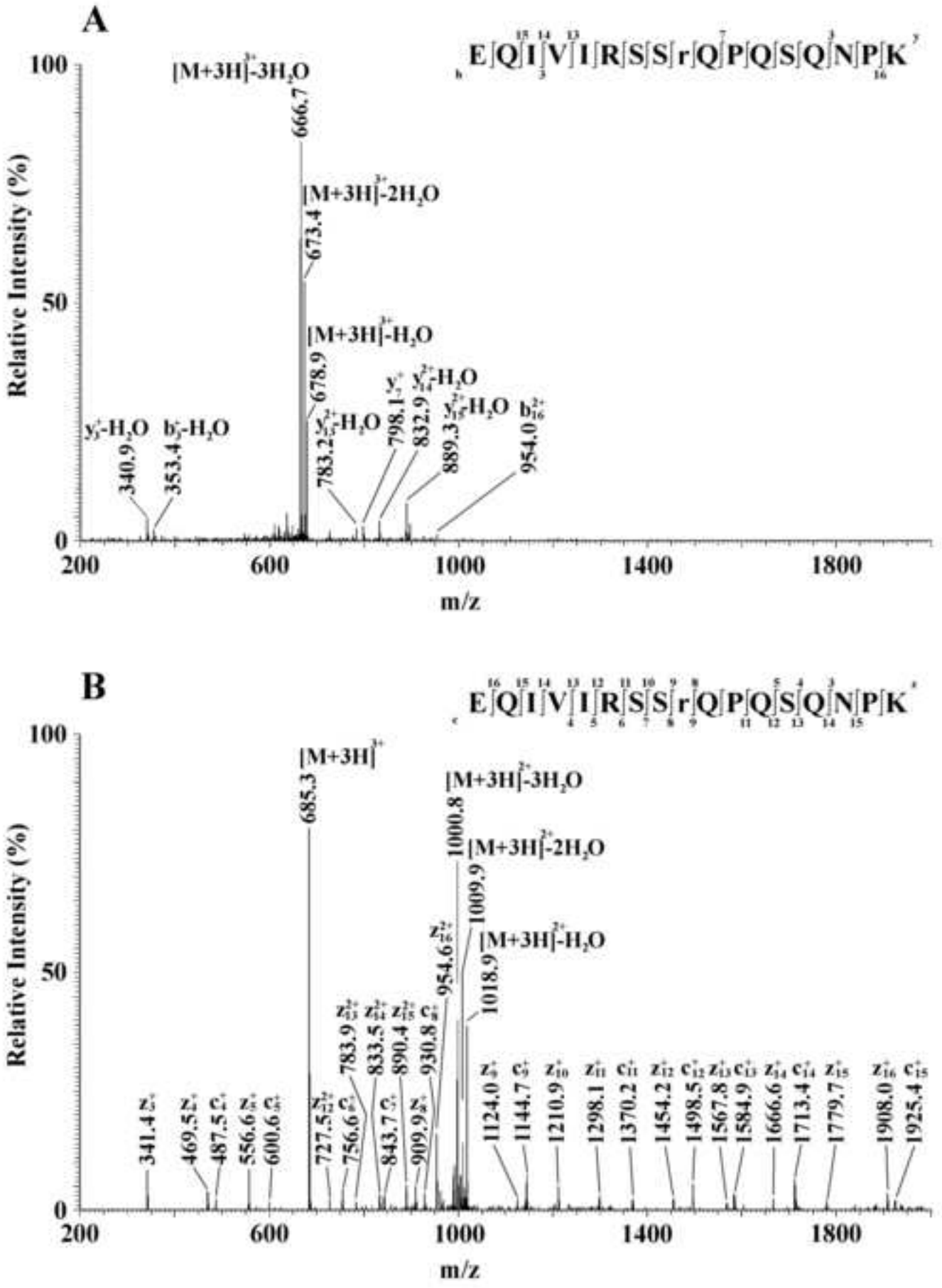


Figure
Click here to download high resolution image



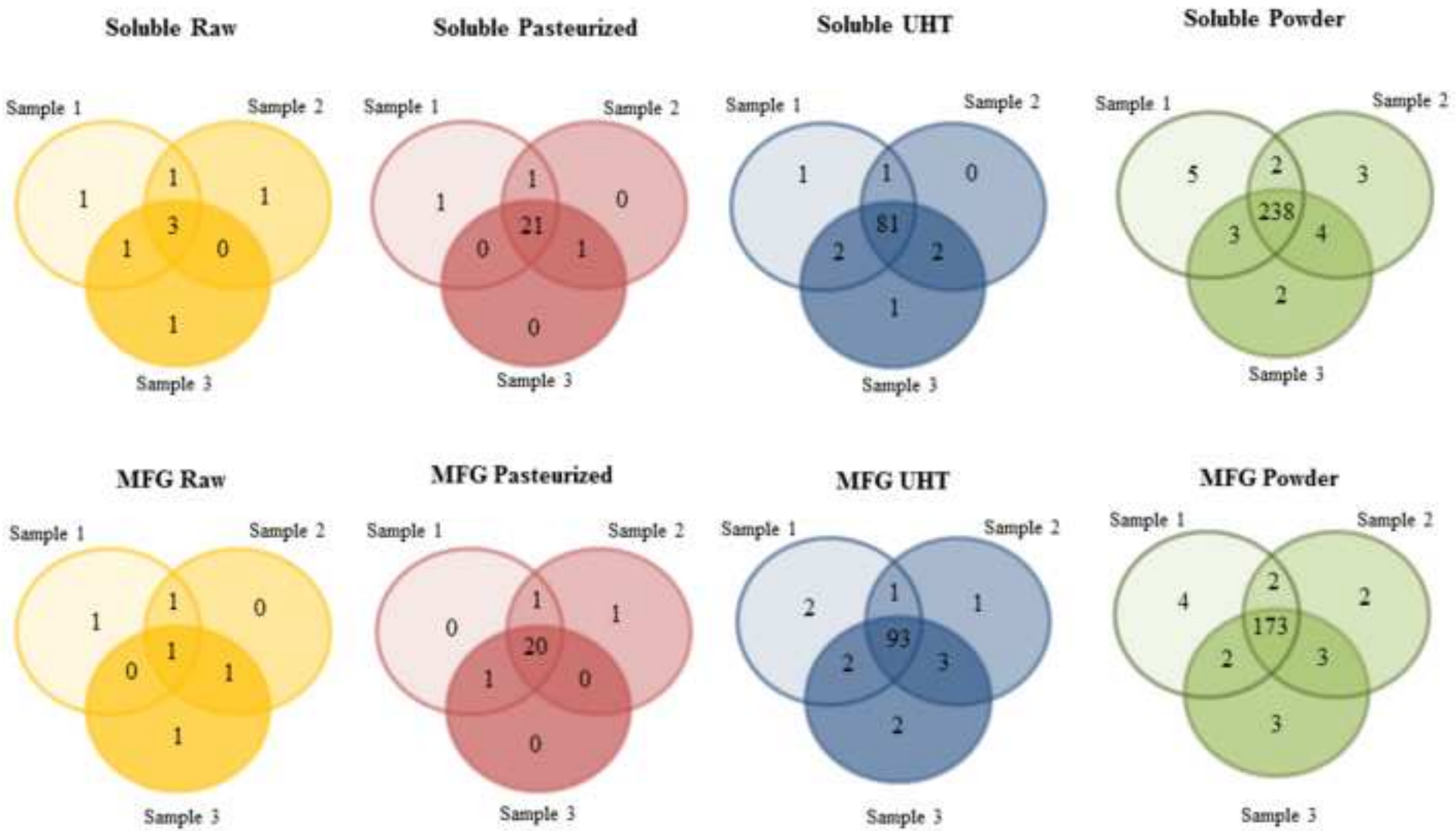
Figure

Click here to download high resolution image



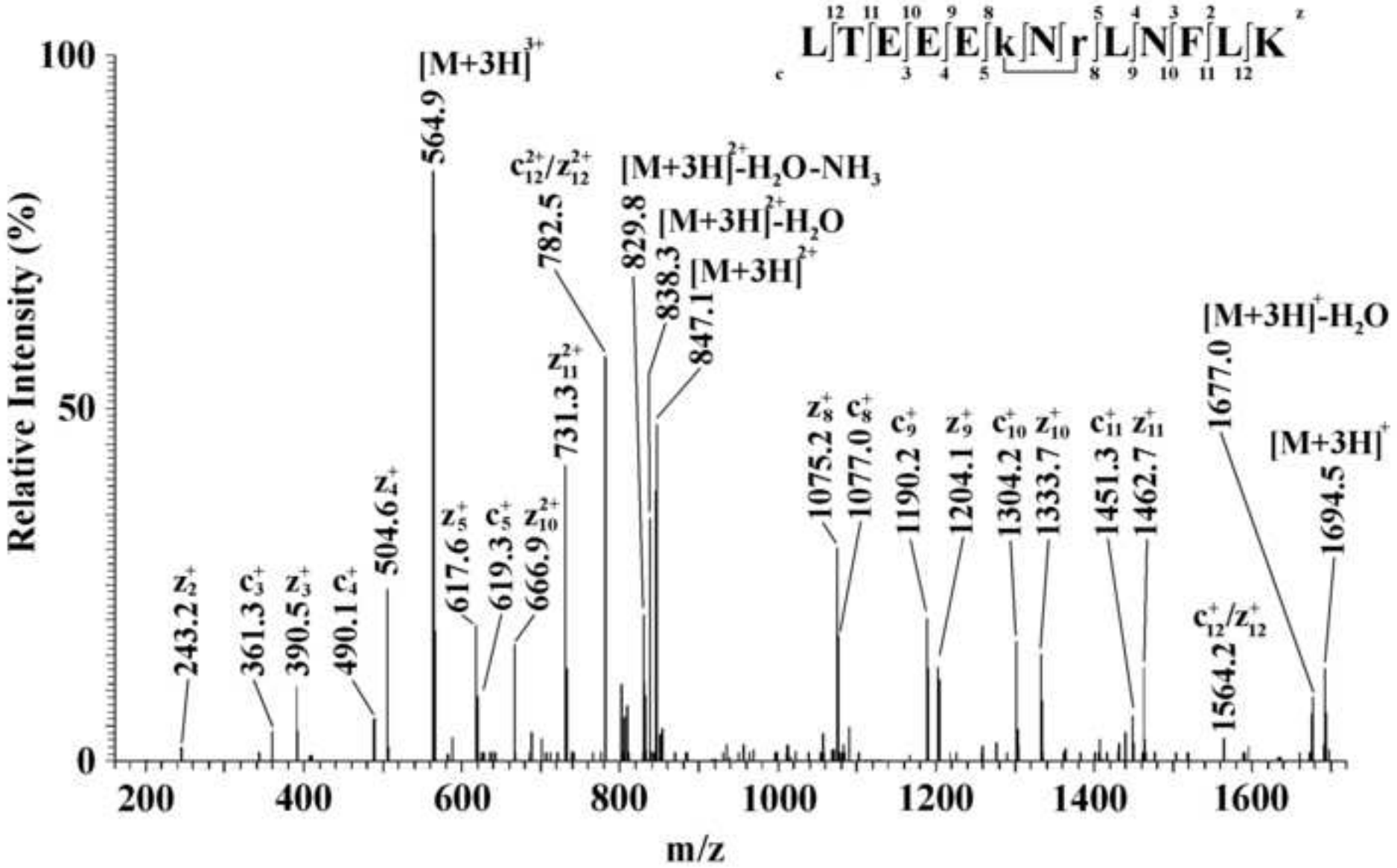
Figure

Click here to download high resolution image



Figure

Click here to download high resolution image



Supplementary material

Click here to download Supplementary material: Renzone_Suppl_Information_revised.docx

Supplementary material

Click here to download Supplementary material: SupplTab2_revised.xlsx

Supplementary material

Click here to download Supplementary material: SupplTab3_revised.xlsx

Supplementary material

Click here to download Supplementary material: SupplTab4_revised.xlsx

Supplementary material

Click here to download Supplementary material: SupplTab5_revised.xlsx

Supplementary material

Click here to download Supplementary material: SupplTab6_revised.xlsx

Supplementary material

Click here to download Supplementary material: SupplTab7_revised.xlsx

Supplementary material

Click here to download Supplementary material: SupplTab8_revised.xlsx

Supplementary material

Click here to download Supplementary material: SupplTab9_revised.xlsx

Supplementary material

Click here to download Supplementary material: SupplFig1.tif

Supplementary material

Click here to download Supplementary material: SupplFig2A.tif

Supplementary material

Click here to download Supplementary material: SupplFig2B.tif

Supplementary material

Click here to download Supplementary material: SupplFig2C.tif

Supplementary material

Click here to download Supplementary material: SupplFig3.tif

Supplementary material

Click here to download Supplementary material: SupplFig4.tif

Supplementary material

Click here to download Supplementary material: SupplFig5_revised.tif

Supplementary material

Click here to download Supplementary material: SupplFig6_revised.tif

***Conflict of Interest**

Click here to download Conflict of Interest: coi_disclosure.pdf