

Glycopeptide Enrichment Using a Combination of ZIC-HILIC and Cotton Wool for Exploring the Glycoproteome of Wheat Flour Albumins

Plaipol Dedvisitsakul,[†] Susanne Jacobsen,[†] Birte Svensson,[†] Jakob Bunkenborg,^{‡,⊥} Christine Finnie,[§] and Per Högglund^{*,†}

[†]Enzyme and Protein Chemistry, Søtofts Plads Building 224, Department of Systems Biology, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark

[‡]Department of Clinical Biochemistry, Copenhagen University Hospital Hvidovre, DK-2650 Hvidovre, Denmark

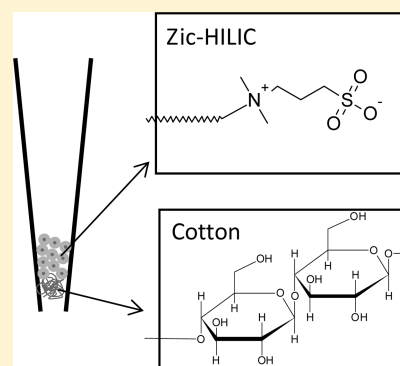
[§]Agricultural and Environmental Proteomics, Søtofts Plads Building 224, Department of Systems Biology, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark

[⊥]Center of Experimental Bioinformatics, Department of Biochemistry and Molecular Biology, University of Southern Denmark, Campusvej 55, DK-5230 Odense, Denmark

Supporting Information

ABSTRACT: Hydrophilic liquid chromatography (HILIC) is used extensively as a sample preparation step for glycopeptide enrichment in proteome research. Here, we have applied cotton wool and a zwitterionic HILIC (ZIC-HILIC) resin in solid-phase extraction microcolumns to provide a higher loading capacity and broader specificity for glycopeptide enrichment. This strategy was applied to tryptic digests of wheat flour albumin extracts followed by simultaneous site-specific ¹⁸O labeling and deglycosylation using peptide-N-glycosidase A (PNGase A) in H₂¹⁸O. Subsequent LC-MS/MS analysis allowed for assignment of 78 N-glycosylation sites in 67 albumin proteins. Bioinformatic analysis revealed that several of the identified glycoproteins show sequence similarity to known food allergens. In addition, the potential impact of some of the identified glycoproteins on wheat beer quality is discussed.

KEYWORDS: Glycopeptide enrichment, HILIC, cotton, N-glycosylation, mass spectrometry, deglycosylation, deamidation, isotope labeling, cereal proteomics



INTRODUCTION

Glycosylation is a major post-translational protein modification known to affect protein folding, stability, activity, molecular interaction, and functionality.¹ N- and O-linked glycosylations represent the most common types of glycosylation.² The majority of proteins secreted into the endoplasmic reticulum (ER) are modified by attachment of an oligosaccharide precursor (Glc₃Man₉GlcNAc₂) onto asparagine residues in N-X-S/T/C consensus sequences,^{3,4} where X can be any amino acid other than proline.⁵ In contrast, there appears to be no strict consensus sequence guiding the attachment of O-linked glycans from soluble nucleotide sugars to the hydroxyl groups of serine or threonine. During transport along the secretory pathway, the N-linked oligosaccharide undergoes a number of maturation steps. First, glucose (Glc) and mannose (Man) residues are removed; subsequently, new sugar residues are added to generate either high-mannose-, hybrid-, or complex-type N-glycans.⁶ For each of these types, a wide variety of glycan structures have been described, and the occupancy of these glycans at different glycosylation sites lead to glycoprotein heterogeneity. Complex-type plant N-glycans differ from those

produced in mammalian glycoproteins. For instance, α (1,6)-fucose (Fuc) and terminal N-acetyl-neuraminic acid (NeuAc) are characteristic of mammals, whereas core β (1,2)-xylose (Xyl) and α (1,3)-Fuc are found in plants.⁶ Currently, very little information is available on the functional importance of glycosylated proteins from plants as compared to mammals. However, it has been reported that glycosylation of plant glycoproteins is required for enzyme activity,^{7,8} thermostability and folding,⁸ oligomerization,⁷ and host–pathogen interaction.⁹

In mass spectrometry, the ionization of glycopeptides is less efficient compared to nonglycosylated peptides, leading to signal suppression and poor detection of glycopeptides.¹⁰ Therefore, enrichment of glycopeptides prior to mass spectrometry is often required. Several approaches for glycopeptide enrichment are available based on, for example, affinity toward glycan-specific lectins^{11,12} and *cis*-diol periodate oxidation followed by hydrazide coupling.^{13,14} Glycopeptides are also frequently separated from more hydrophobic non-

Received: December 20, 2013

glycosylated peptides by hydrophilic interaction liquid chromatography (HILIC). Several HILIC materials have been used for glycopeptide enrichment, such as zwitterionic (ZIC),^{15–18} amide-based (e.g., TSKgel Amide-80),¹⁸ and carbohydrate-based HILIC material (e.g., Sepharose, cellulose, and cotton wool).^{19,20} The retention of glycans in ZIC-HILIC involves polar and ionic interactions,²¹ but with carbohydrate-based materials such as cotton, the retention is mainly based on hydrogen bonding.²⁰ These HILIC materials may thus provide distinctly different glycopeptide profiles. In addition to the glycan composition, the amino acid composition also determines the hydrophilicity of glycopeptides. Thus ion-pairing agents that neutralize charges on amino acid residues have been demonstrated to improve glycopeptide enrichment by ZIC-HILIC.²²

Compared to other HILIC materials, ZIC-HILIC shows a higher selectivity for glycopeptides.¹⁸ In addition to N- and O-linked glycopeptides, ZIC-HILIC has also been applied successfully to enrich glycosylphosphatidylinositol-anchored glycopeptides.²³ For proteomic analysis, the ZIC-HILIC material is typically prepared in a custom-made microcolumn packed in a pipet tip.^{16,24} Construction of these columns is, however, technically challenging because of the small particle size (10 μm) and because the loading capacity is limited. One way to solve the capacity problem is to increase the amount of ZIC-HILIC resin in the microcolumn, but flow velocity is very low as a result of the high backpressure. HILIC columns based on cotton wool, however, are cheap, very easy to prepare, and also allow fast flow.²⁰

In the present work, we implemented a combination of cotton wool and ZIC-HILIC (referred to as ZIC-cotton HILIC) for larger-scale glycopeptide enrichment in a microcolumn format compatible with sensitive mass spectrometric analysis. ZIC-cotton was applied to enrich tryptic N-glycopeptides prepared from wheat flour albumin proteins followed by deglycosylation and LC–MS/MS analysis for glycosylation site mapping. To date, only a few wheat glycoproteins have been identified, and in most cases, details on their molecular structures remain to be described. In addition, most information was obtained from studies of single purified glycoproteins, and the global glycosylation profile of the wheat proteome has not yet been described. The glycoproteins and glycosylation sites identified in the present study thus contribute to a more comprehensive description of wheat flour components and provide a basis for future research on the biological function of plant proteins and food glycoprotein allergens. The long-term potential benefits include discovering novel targets for improvement of wheat-based food products (e.g., bread and beer) and so-called molecular pharming (i.e., recombinant production of high-value therapeutic proteins in cereals).

MATERIALS AND METHODS

Protein Samples

Wheat albumin proteins were prepared as previously described²⁵ with some modifications. Briefly, 100 mg of grain powder was mixed with 25 mM sodium phosphate buffer, pH 7.5 (1 mL), at 4 °C for 60 min, insoluble debris was removed by centrifugation, and the supernatant was precipitated by an equal volume of 20% (w/v) TCA. Horseradish peroxidase (HRP) and bovine serum albumin (BSA) were purchased from Sigma-Aldrich (St. Louis, MO).

In-Solution Trypsin Digestion

Protein samples (approximately 40 μg) were redissolved in 400 mM NH_4CO_3 and 8 M urea (21 μL) and reduced by addition of 5 μL of 45 mM DTT. The samples were then incubated at 50 °C for 20 min and thereafter chilled on ice. Cysteine alkylation was performed by adding 100 mM iodoacetamide (5 μL) and incubating at room temperature (20–23 °C for 45 min). Lastly, 140 μL of H_2O and 5 μL (0.04 $\mu\text{g}/\text{mL}$) of sequencing-grade modified trypsin (Promega, Madison, WI) was added followed by incubation at 37 °C for 24 h. Samples were subsequently stored at –20 °C until use.

Glycopeptide Enrichment

ZIC-HILIC and cotton wool microcolumns were essentially prepared as described.^{16,20} Cotton wool microcolumns were made by pushing a small piece of cotton wool taken from a cotton wool pad into a 20 μL GELoader tip (Eppendorf, Hamburg, Germany). For ZIC-cotton HILIC microcolumns, the ZIC-HILIC resin (a kind gift of Sequant, Umeå, Sweden; particle size 10 μm) was packed on top of the cotton wool (Supporting Information Figure 1). In-solution trypsin-digested protein samples (20 μg) were dried (SPD 1010 Speed Vac, Thermo Scientific), redissolved in either (i) 80% (v/v) acetonitrile (ACN) and 0.1% of trifluoroacetic acid (TFA) or (ii) 80% (v/v) acetonitrile (ACN) and 0.5% (v/v) formic acid (FA), and applied onto HILIC microcolumns equilibrated with the same solvents. The column was subsequently washed twice followed by the elution of bound peptides in 40 μL of 99.5% H_2O and 0.5% FA.

Deglycosylation by PNGase A and ^{18}O Labeling

Glycopeptide eluates from ZIC-cotton HILIC microcolumns were dried (SPD 1010 Speed Vac) and redissolved in 100 mM ammonium acetate (pH 5.5) prepared in H_2^{18}O . Deglycosylation was performed for 24 h at 37 °C with 0.1 mU (enzyme activity that hydrolyzes 0.1 μM ovalbumin glycopeptide within 1 min at 37 °C) PNGase A (Roche Applied Science, Mannheim, Germany). Subsequently, deglycosylated samples were dried (SPD 1010 Speed Vac) and redissolved in 100 mM ammonium acetate (pH 5.5) in normal water (H_2^{16}O). Sequencing-grade modified trypsin (0.2 μg) (Promega, Madison, WI) was added to the samples, and the mixture was incubated as above for 24 h at 37 °C.

MALDI-TOF MS

Samples (1 μL out of 40 μL , which corresponds to an estimated maximum of 0.5 μg total protein) eluted from ZIC-HILIC and cotton wool microcolumns as described above were spotted directly onto a ground steel MALDI target plate (Bruker Daltonics, Bremen Germany), allowed to dry, and overlaid with 1 μL of dihydroxybenzoic acid (DHB) matrix (20 mg/mL in 99.5% H_2O and 0.5% FA). All other samples were purified on R2 reversed-phase microcolumns (loaded/washed in 0.1% TFA and eluted in 80% ACN and 0.1% TFA) and subsequently spotted on the target plate as above. MALDI-TOF MS analyses were performed by Ultraflex II MALDI-TOF/TOF MS (Bruker Daltonics). The obtained mass spectra were processed with FlexAnalysis software (Bruker Daltonics).

LC–MS/MS and Database Search

Peptide samples were desalted with StageTip C18 reversed-phase columns (Thermo Fisher Scientific) and analyzed on an EASY nLC 1000 chromatograph (Thermo Fisher Scientific) with an EASY-Spray column (Pepmap 3 μm , C18 15 cm $\times$ 75 μm) coupled online to a Q Exactive Orbitrap (Thermo Fisher

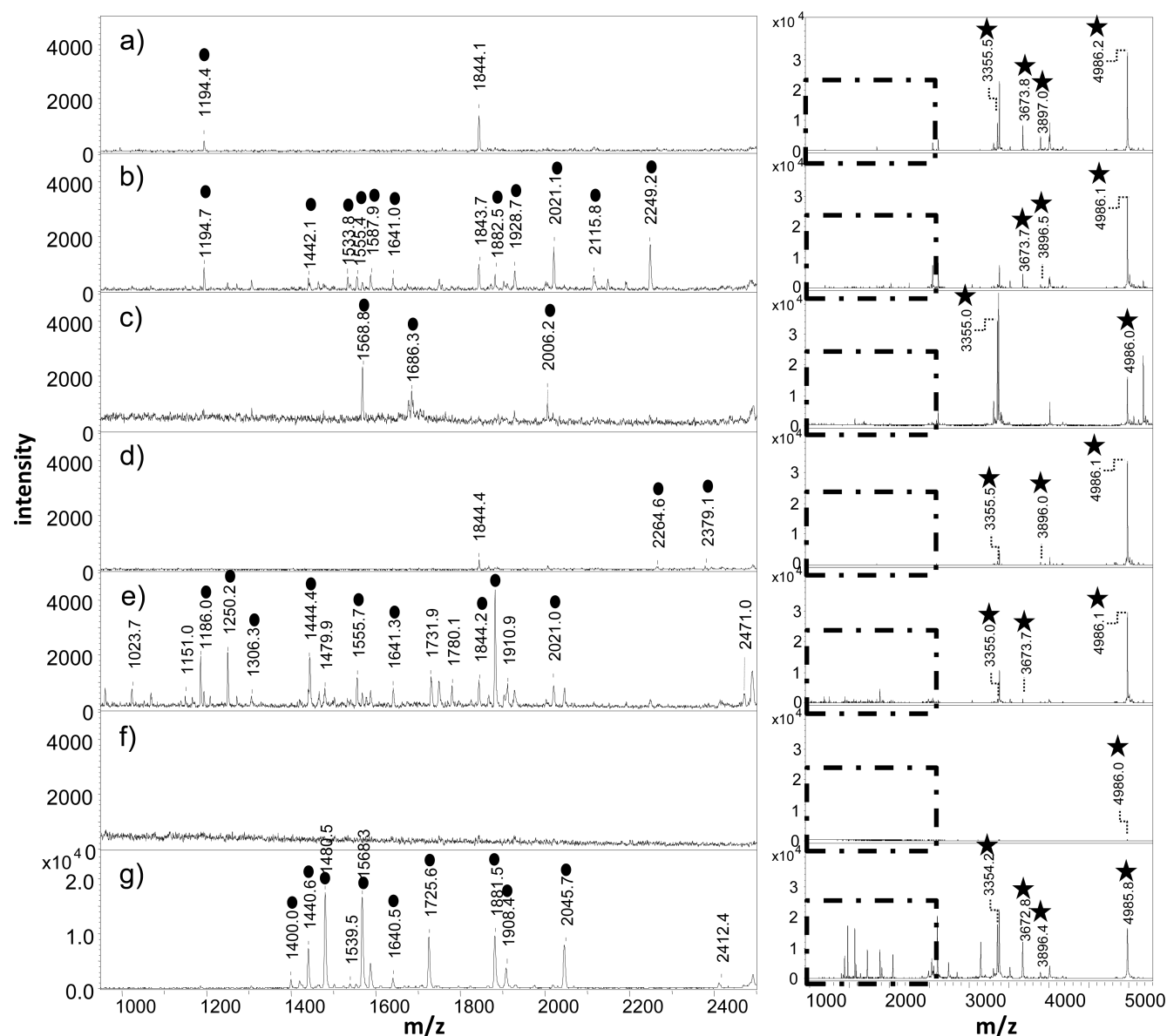


Figure 1. MALDI-TOF MS spectra of HRP-BSA tryptic peptides eluted from microcolumns packed with different resins. Samples were loaded on microcolumns packed with ZIC-cotton HILIC (a, d), cotton wool (b, e), or ZIC-HILIC (c, f) in a mobile phase of either 80% ACN and 0.5% FA (a–c) or 80% ACN and 0.1% TFA (d–f) and were eluted in 99.5% H₂O and 0.5% FA. For the R2 reversed-phase microcolumn (g), the sample was loaded in 0.1% TFA and eluted in 80% ACN and 0.1% TFA. The left panel shows the magnified spectra of peptides from *m/z* 950 to 2500. Ovals indicate signals matching theoretical masses of nonglycosylated tryptic HRP and BSA peptides listed in Supporting Information Tables 1 and 2. Stars indicate signals matching previously determined HRP glycopeptides.²⁶ The mass spectra represent the average of >400 laser shots.

Scientific). The separation was performed using a flow rate of 300 nL min^{−1} and a gradient of solvents A (0.1% FA) and B (80% ACN and 0.1% FA). Raw data acquired in the positive ion mode were converted into the Mascot generic format (MGF) by Proteome Discoverer 1.3 (Thermo Fisher Scientific). Data were searched against HarvEST35 (<http://www.harvest-web.org/>) and TIGR wheat (<http://compbio.dfci.harvard.edu/tgi/>) EST databases using MASCOT (Version 2.2.04, Matrix Science): fixed modifications, carbamidomethyl cysteine; variable modifications, oxidation of methionine and deamidation of Asn with incorporation of ¹⁶O and ¹⁸O; and enzyme, trypsin (1 missed cleavage). MS and MS/MS accuracy were set to 10 ppm and 0.02 Da, respectively. Only glycopeptides containing the consensus tripeptide (N-X-S/T/C) in which the deamidated N was tagged with ¹⁸O atoms were

accepted. The identified EST sequences were annotated by BLAST searches against NCBI (<http://www.ncbi.nlm.nih.gov>). In addition, the identified protein sequences from the HarvEST35 and TIGR wheat databases were searched against the Structural Database of Allergenic Proteins (SDAP; <http://fermi.utmb.edu/SDAP/>) for bioinformatics-based allergen predictions.

RESULTS AND DISCUSSION

Enrichment of Glycopeptides by HILIC

Microcolumns containing reversed-phase or ZIC-HILIC resins, cotton wool, or a combination of ZIC-HILIC resin and cotton wool (ZIC-cotton HILIC) were constructed, and the efficiency for enrichment of glycopeptides was tested by applying a model protein mixture of horseradish peroxidase with nine potential

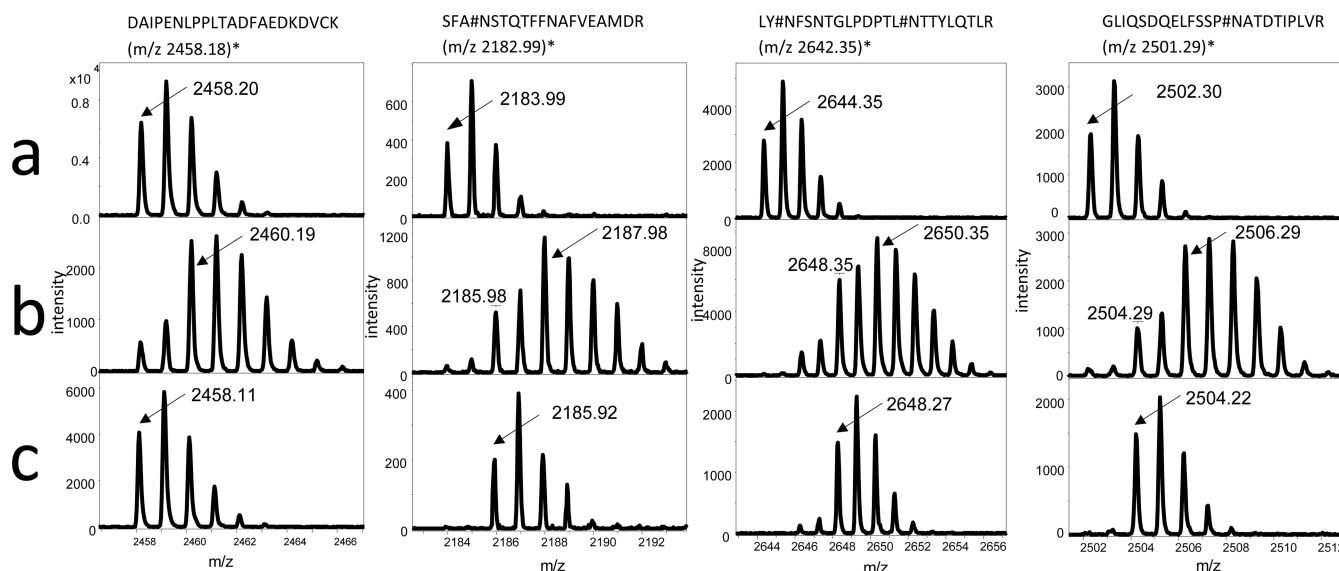


Figure 2. MALDI-TOF MS spectra of selected HRP-BSA tryptic peptides subjected to PNGase A deglycosylation and $^{18}\text{O} \rightarrow ^{16}\text{O}$ back exchange. Deglycosylation with PNGase A was performed in ^{16}O water (a), ^{18}O water (b), or ^{18}O water followed by $^{18}\text{O} \rightarrow ^{16}\text{O}$ C-terminal back exchange (c). The MS spectra were matched to nonglycosylated DAIPENLPPLTADFAEDKDVCK (m/z 2458.18*) and the predicted glycopeptides SFA#NSTQTFFNAFVEAMDR (m/z 2182.99*), LY#NFSNTGLPDPTLN#TTYLQTLR (m/z 2642.35*), and GLIQSDQELFSSPN#ATDTIPLVR (m/z 2501.29*). The # symbol indicates glycosylation sites. The * symbol indicates m/z values calculated with nondeamidated asparagines. The mass spectra represent the average of 1000 laser shots.

N-glycosylation sites and nonglycosylated BSA (HRP-BSA; Supporting Information Tables 1 and 2). Peptides from tryptic digests of HRP-BSA were thus loaded on microcolumns packed with the various resins and analyzed by MALDI-TOF MS (Figure 1). In the sample eluted from a reversed-phase microcolumn, a range of high-intensity signals from HRP-BSA peptides matched predicted nonglycosylated peptides (Figure 1g). In contrast, the mass spectra of samples eluted from HILIC columns (Figure 1a–f) are, in general, dominated by signals in a higher m/z range (>3000). Several of these match previously reported HRP glycopeptides.²⁶ Although weak signals of a considerable number of nonglycosylated peptides were observed after enrichment by cotton wool (Figure 1b,e), most of these were absent in the mass spectra from ZIC-cotton (Figure 1a,d). This indicates that cotton wool supplemented with ZIC-HILIC resin shows increased selectivity for glycopeptides. With ZIC-HILIC, only two glycopeptides were detected (Figure 1c). This result possibly reflects either a more narrow specificity of ZIC-HILIC compared to cotton wool or a lower capacity of ZIC-HILIC because of the limited amount of resin in the microcolumn. It is also plausible that glycopeptides are bound to the ZIC-HILIC resin but are not efficiently eluted under the conditions used.

The use of TFA (Figure 1d–f) instead of FA (Figure 1a–c) in the sample loading and washing solvents did not efficiently reduce the binding of nonglycosylated peptides. In an acidic mobile phase, the negatively charged TFA ion pairs with positively charged amino acids, thereby reducing both hydrophilic and weak electrostatic interactions of nonglycosylated peptides with the ZIC-HILIC resin.²² Cotton wool is, however, uncharged, and nonglycosylated peptides are expected to be retained mainly via hydrogen-bond interactions that are not blocked by ion pairing with TFA. The coenriched nonglycosylated peptides ETYGDMAECCEKQEPER (m/z 2021.1), VHKECCHGDLLECADDR (m/z 2115.3), FKDLGEEHFK (m/z 1250.4), YICDNQDTISSK (m/z 1444.5),

RPCFSALTPDETYVPK (m/z 1882.1), and DDPHACYSTVFDK (m/z 1555.7) contain side-chain carboxylic acids (aspartate and glutamate), hydroxyls (serine and threonine), and amides (carboxyamidomethylated cysteine) that might hydrogen bond with the abundant hydroxyl groups in the cellulose.

Deglycosylation of Glycopeptides and ^{18}O Labeling

HRP-BSA peptides eluted from HILIC microcolumns were subjected to deglycosylation using PNGase A, which hydrolyzes the bond between asparagine and the innermost GlcNAc of N-glycans. In contrast to the commonly used PNGase F, PNGase A is also active on N-glycans with α -1,3-linked core fucosylation that is common in glycoproteins from insects²⁷ and plants.⁶ In this process, the asparagine residue is converted to aspartate through deamidation. Glycosylation sites can thus be determined through the resulting net mass increase of 0.98 Da. In order to distinguish the glycosylation sites from other deamidated asparagines, deglycosylation was performed in H_2^{18}O , resulting in a unique mass increment of 2.99 Da for glycosylated residues subjected to the reaction with PNGase A that may be used in database search programs.^{28,29}

When the mass spectra of peptides in HRP-BSA tryptic digests subjected to PNGase A in H_2^{18}O and H_2^{16}O were compared, major alterations in the isotopic envelope profiles were observed, and for nonglycosylated peptides, the monoisotopic signals were increased by approximately 2 Da, indicating that ^{18}O labeling was not specific to glycosylation sites (Figure 2a, b). This is probably due to partial incorporation of ^{18}O in the C-terminal carboxylate of the peptide backbone mediated through exchange catalyzed by residual trace amounts of trypsin as previously reported.³⁰ Labeling of C-terminal carboxyl groups complicates data analysis and may result in false positive identification of glycosylation sites.³¹ To minimize ^{18}O labeling of the C-terminus, we thus subjected the ^{18}O -labeled peptides to a second round of carboxyl-oxygen exchange activity with trypsin

in the presence of H_2^{16}O . After this exchange reaction, no significant mass shifts were observed for nonglycosylated peptides subjected to PNGase A in H_2^{18}O and H_2^{16}O , whereas peptides with single and double N-glycosylation sites displayed peak shifts of approximately +2 Da and +4 Da, respectively, indicative of successful incorporation of ^{18}O at the deamidated asparagines and depletion from the C-termini (Figure 2c). Attempts to minimize ^{18}O labeling at the C-termini by using a protease inhibitor (Pefabloc) and heat inactivation of trypsin were less efficient than the trypsin-catalyzed back exchange method (Supporting Information Figure 2).

Global Analysis of N-Glycosylation Sites in Wheat Protein Extracts

On the basis of the results from the model system with HRP-BSA, a strategy was devised for global screening of N-glycosylation sites in proteins from the albumin fraction of wheat flour (Figure 3). First, cotton wool and ZIC-cotton

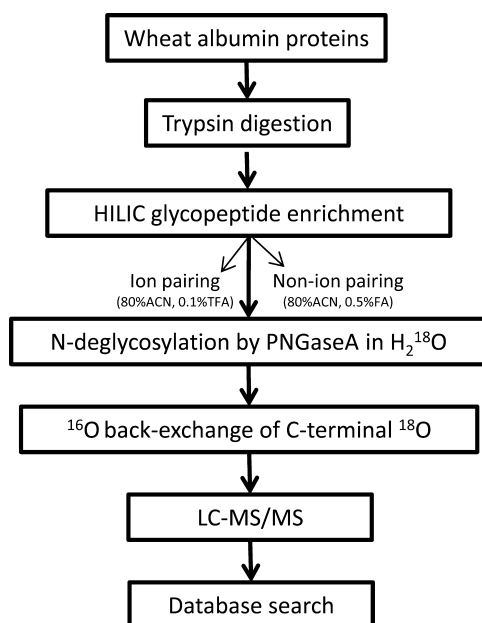


Figure 3. Overview of the strategy to screen for N-glycosylation sites in wheat protein extracts.

microcolumns were employed to enrich glycopeptides from tryptic digests using mobile phases containing either 0.1% TFA or 0.5% FA in 80% ACN. To enable preliminary assessment of the enrichment efficiency, samples of eluted wheat albumin tryptic peptides were analyzed by MALDI-TOF MS (Supporting Information Figure 3). Signals at higher m/z ratios are seen in samples enriched by cotton wool and ZIC-cotton HILIC both in ion-pairing (80% ACN and 0.1% TFA) and non-ion-pairing modes (80% ACN and 0.5% FA) compared to a sample eluted from a R2 reversed-phase microcolumn. Different MALDI-TOF MS spectra patterns were observed for peptides enriched by cotton wool and ZIC-cotton both in ion-pairing and non-ion-pairing modes, suggesting that different peptide profiles are obtained by using these types of HILIC stationary phases.

The peptides eluted from the cotton wool and ZIC-cotton microcolumns were subjected to digestion with PNGase A in the presence of H_2^{18}O followed by trypsin-mediated back exchange in H_2^{16}O as described above, and the deglycosylated peptides were subject to LC-MS/MS analysis on a Q-Exactive

Orbitrap. In this way, we identified 78 different glycosylation sites assigned to 67 albumin glycoproteins (Supporting Information Table 3). In all identified glycosylation sites, asparagine residues in NXS/T/C sequons were deamidated and labeled with ^{18}O . An MS/MS fragment ion spectrum of the deglycosylated peptide LDMDVTLV#NITR matched to TC246874 in the TIGR wheat database is shown as an example in Figure 4 (MS/MS spectra for all identified peptides are displayed in Supporting Information Table 4).

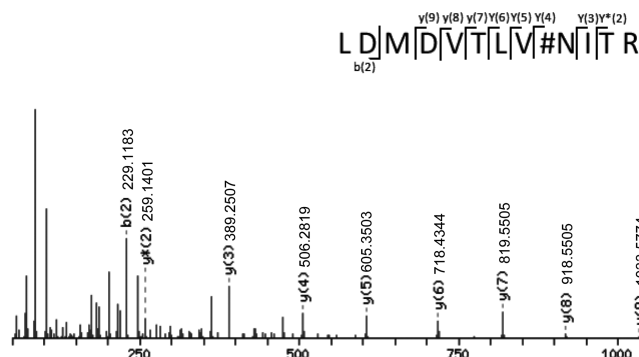


Figure 4. Fragment ion spectrum of deglycosylated peptide LDMDVTLV#NITR matched to TC246874 in the TIGR wheat database. The # symbol indicates that the N-linked asparagine was converted to ^{18}O -incorporated aspartic acid after deglycosylation with PNGase A. The * symbol indicates loss of ammonia.

For both ion-pairing and non-ion-pairing modes, a higher number of glycopeptides were identified by ZIC-cotton as compared to cotton wool (Figure 5), presumably reflecting a broader glycopeptide selectivity of ZIC-cotton. A limited binding capacity of cotton wool was also considered, but the number of glycopeptides identified did not increase when an approximately 3-fold larger amount of cotton was used (data not shown). With both ZIC-cotton and cotton wool, fewer glycopeptides were identified when using TFA as the ion-pairing agent compared to the non-ion-pairing mode using FA (Figure 5).

According to our survey of the dbPTM³² and UniProt databases, most glycosylation sites identified here have not

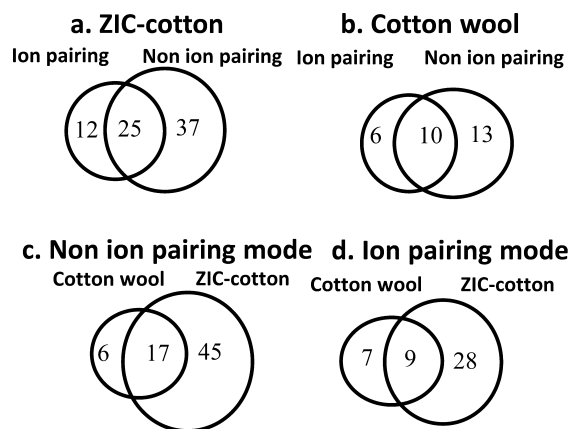


Figure 5. Venn diagram summarizing the number of glycosylation sites identified after glycopeptide enrichment by ZIC-cotton and cotton wool in ion-pairing (mobile phase 80% ACN and 0.1% TFA) and non-ion-pairing modes (mobile phase 80% ACN and 0.5% FA). In total, 78 different glycosylation sites were assigned to 67 albumin glycoproteins.

previously been experimentally verified; the data provided here thus provides novel insight into the composition of cereal proteomes that may be beneficial for further studies to gain deeper understanding on the structure and biological roles of individual proteins, as discussed below.

Potential Allergenicity of Wheat Glycoproteins

Major food allergens are glycoproteins,³³ and most IgE-mediated allergic reactions involve water-soluble proteins.³⁴ Generally, an extensive immune cross-reactivity is found among different plant and insect species, and the allergenicity of plant glycoproteins is related to the presence of xylose and α -1,3-linked core fucose residues that do not exist in mammalian glycoproteins.^{35,36} Thus, potential problems with IgE cross-reactivity may be avoided by engineering the plant glycoprotein processing machinery, for example, for production of biopharmaceutical proteins.³⁷

Recently, Palmisano et al. applied the Structural Database of Allergenic Proteins (SDAP; <http://fermi.utmb.edu/SDAP/>) to screen identified grape glycoproteins for similarity to known allergens.³⁸ In the present study, the same methodology revealed that 21 of the total 67 identified glycoproteins from wheat albumins have amino acid sequence similarity to food allergens (Supporting Information Table 3). Among these, dimeric α -amylase inhibitor (TC249880), 27K protein (TC250399), and serpin-N3.2 (TC265635) are similar to well-known wheat allergens called Tri a TAI, Tri a 27.0101, and Tri a 33.0101, respectively.^{39,40} α -Amylase inhibitors and serpins have been identified as major allergens that cause baker's asthma.⁴¹ Thaumatin-like protein (U35_14143) and xylanase inhibitor protein-1 (TC264929, TC264927) show similarity to allergens Mus a 4.0101 and Ziz m 1.0101, respectively. These proteins were recently identified as wheat allergens by immunoblotting with IgE from an atopic dermatitis patient.⁴² Wheat β -D-glucan exohydrolase (TC250803), triticin (TC264477), and globulin-3A (TC234045), also reported to bind IgE,⁴³ are homologues to Asp n 14, Pis v 2, and Ses i 3, respectively. Fructan 1-exohydrolase (TC253637), globulin-1S (TC246874), polygalacturonase-like protein (TC267315, U35_5393), and neutral ceramidase (U35_4630) show similarity to the allergens Lyc e 2.0102, Cor a 11, Phl p 13, and Pha a 5, respectively. Because cross-reactivity among different plant species can be expected, these glycoproteins are obvious targets in the identification of novel wheat allergens.

Wheat β -glucosidase does not have similarity to any allergenic proteins in the SDAP database but was reported to bind IgE.^{42,43} This underscores that identification of wheat allergens is still at a preliminary stage and that allergenic proteins may not yet be available in the database. A 36 kDa allergen causing baker's asthma was identified as seed-specific peroxidase from wheat flour,⁴⁴ and wheat peroxidase 1 has also been identified as an allergen.^{42,45} Six peroxidases were identified as glycoproteins in the present study but did not match allergens in SDAP. However, a barley homologue (gil 326513532) of the identified putative peroxidase (U35_22623) has similarity to the Tri a gliadin allergen. These data provide a basis for further study of the differences among allergenic and nonallergenic isoforms of cereal peroxidases.

The proteins identified here may have potential as natural immunomodulatory glycoproteins to enhance immune responses. One of the most useful approaches to modulate the immune system has been to identify natural compounds, such as phyto glycoproteins,⁴⁶ because of their immunogenic

carbohydrate epitopes. In order to evaluate the allergenic potential of the identified glycoprotein targets, further analysis of possible IgE binding via carbohydrate moieties on glycoproteins can be performed, for example, by preincubation of the candidate allergenic glycoproteins with IgE from patient serum and testing for binding of polyclonal antihorseradish peroxidase antibodies, which recognize N-linked glycan moieties with an α (1,3)-fucose branch.⁴⁷

Glycoproteins Related to Beer Production

Many of the proteins present in beer are derived from cereal grains. Understanding the effects of these proteins on beer quality (e.g., foam and haze formation) requires information about characteristics and composition. Glycated and glycosylated proteins are reported to influence the quality of beer,⁴⁸ but only a few studies have reported on cereal glycoproteins.⁴⁸ In addition, allergenicity in beer is also an issue of beer quality, and glycoprotein identification may thus contribute to the evaluation of beer quality.

Many proteins influencing beer quality have been identified by proteomic studies of the beer,^{49,50} beer foam,^{51–53} and beer haze.⁵⁴ Most information about the influence of beer proteins has been reported for barley beers, whereas much less is available on wheat beer,^{49,50} and, to our knowledge, no protein profiling has been reported for wheat beer. Pathogenesis-related proteins (PRs) are considered to be the most important proteins for beer production because of their thermal stability and resistance to proteolysis as well as because of their role in foam and haze formation.⁵⁵ In the present study, we identified several PRs as glycoproteins (Supporting Information Table 3), including dimeric amylase inhibitor (TC249880), serpin (TC265635), peroxidases (TC249600, TC251690, TC240657, U35_22623, and U35_14350), and thaumatin-like protein (U35_14143).

CONCLUSIONS

The combination of ZIC-HILIC and cotton wool provides a marked improvement in the use of HILIC-based stationary phases for the enrichment of glycopeptides from complex protein mixtures. This is, to our knowledge, one of the first examples where HILIC together with enzymatic deglycosylation in $H_2^{18}O$ has been applied to large-scale glycopeptide identification in plants.

ASSOCIATED CONTENT

Supporting Information

Predicted tryptic peptides of horseradish peroxidase; predicted tryptic peptides of bovine serum albumin; N-glycoproteins and their glycosylation sites identified in wheat albumin proteins; MS/MS spectra of identified glycopeptides from wheat albumin showing the fragment ion series of the deglycosylated peptides labeled with ^{18}O ; representative images of custom-made HILIC microcolumns; comparison of approaches to avoid ^{18}O labeling at the C-terminal carboxyl group by using a protease inhibitor (Pefabloc), heat inactivation, and trypsin-catalyzed back exchange; and MALDI-TOF MS spectra of tryptic peptides from wheat albumin proteins eluted from ZIC-cotton and cotton wool microcolumns. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*Tel: +45 4525 5503; Fax: +45 4588 6307; E-mail: ph@bio.dtu.dk.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Birgit Andersen is acknowledged for technical assistance with mass spectrometric analysis and Louise Helsted and Julie Henriksen, for providing the wheat albumin sample. We are grateful to Sequant (Umeå, Sweden) for providing ZIC-HILIC resin material. The Danish Council for Independent Research | Natural Sciences and the Centre for Advanced Food Studies are acknowledged for funding Q-Exactive Orbitrap MS and Ultraflex MALDI-TOF MS equipment, respectively. Plaipol Dedvisitsakul is supported by a Ph.D. scholarship from the Thai Royal Government.

REFERENCES

- (1) Spiro, R. G. Protein glycosylation: Nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology* **2002**, *12*, 43R–56R.
- (2) Fitchette, A. C.; Dinh, O. T.; Faye, L.; Bardor, M. Plant proteomics and glycosylation. *Methods Mol. Biol.* **2007**, *355*, 317–342.
- (3) Bause, E.; Hettkamp, H. Primary structural requirements for N-glycosylation of peptides in rat liver. *FEBS Lett.* **1979**, *108*, 341–344.
- (4) Bause, E.; Legler, G. The role of the hydroxy amino acid in the triplet sequence Asn-Xaa-Thr (Ser) for the N-glycosylation step during glycoprotein biosynthesis. *Biochem. J.* **1981**, *195*, 639–644.
- (5) Helenius, A.; Aebi, M. Roles of N-linked glycans in the endoplasmic reticulum. *Annu. Rev. Biochem.* **2004**, *73*, 1019–1049.
- (6) Song, W.; Henquet, M. G. L.; Mentink, R. A.; van Dijk, A. J.; Cordewener, J. H. G.; Bosch, D.; America, A. H. P.; van der Krol, A. R. N-Glycoproteomics in plants: Perspectives and challenges. *J. Proteomics* **2011**, *74*, 1463–1474.
- (7) Kimura, Y.; Hess, D.; Sturm, A. The N-glycans of jack bean alpha-mannosidase. Structure, topology and function. *Eur. J. Biochem.* **1999**, *264*, 168–175.
- (8) Lige, B.; Shengwu, M.; van Huystee, R. The effects of the site-directed removal of N-glycosylation from cationic peanut peroxidase on its function. *Arch. Biochem. Biophys.* **2001**, *386*, 17–24.
- (9) Lim, J. M.; Aoki, K.; Angel, P.; Garrison, D.; King, D.; Tiemeyer, M.; Bergmann, C.; Wells, L. Mapping glycans onto specific N-linked glycosylation sites of *Pyrus communis* PGIP redefines the interface for EPG-PGIP interactions. *J. Proteome Res.* **2009**, *8*, 673–680.
- (10) Geyer, H.; Geyer, R. Strategies for analysis of glycoprotein glycosylation. *Biochim. Biophys. Acta* **2006**, *1764*, 1853–1869.
- (11) Minic, Z.; Jamet, E.; Négroni, L.; der Garabedian, P. A.; Zivy, M.; Jouanin, L. A sub-proteome of *Arabidopsis thaliana* mature stems trapped on Concanavalin A is enriched in cell wall glycoside hydrolases. *J. Exp. Bot.* **2007**, *58*, 2503–2512.
- (12) Catala, C.; Howe, K. J.; Hucko, S.; Rose, J. K.; Thannhauser, T. W. Towards characterization of the glycoproteome of tomato (*Solanum lycopersicum*) fruit using Concanavalin A lectin affinity chromatography and LC-MALDI-MS/MS analysis. *Proteomics* **2011**, *11*, 1530–1544.
- (13) Zhang, H.; Li, X. J.; Martin, D. B.; Aebersold, R. Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat. Biotechnol.* **2003**, *21*, 660–666.
- (14) Zhou, Y.; Aebersold, R.; Zhang, H. Isolation of N-linked glycopeptides from plasma. *Anal. Chem.* **2007**, *79*, 5826–5837.
- (15) Picariello, G.; Ferranti, P.; Mamone, G.; Roepstorff, P.; Addeo, F. Identification of N-linked glycoproteins in human milk by hydrophilic interaction liquid chromatography and mass spectrometry. *Proteomics* **2008**, *8*, 3833–3847.
- (16) Häggglund, P.; Bunkenborg, J.; Elortz, F.; Jensen, O. N.; Roepstorff, P. A new strategy for identification of N-glycosylated proteins and unambiguous assignment of their glycosylation sites using HILIC enrichment and partial deglycosylation. *J. Proteome Res.* **2004**, *3*, 556–566.
- (17) Häggglund, P.; Matthiesen, R.; Elortz, F.; Højrup, P.; Roepstorff, P.; Jensen, O. N.; Bunkenborg, J. An enzymatic deglycosylation scheme enabling identification of core fucosylated N-glycans and O-glycosylation site mapping of human plasma proteins. *J. Proteome Res.* **2007**, *6*, 3021–3031.
- (18) Wohlgemuth, J.; Karas, M.; Eichhorn, T.; Hendriks, R.; Andrecht, S. Quantitative site-specific analysis of protein glycosylation by LC-MS using different glycopeptide-enrichment strategies. *Anal. Biochem.* **2009**, *395*, 178–188.
- (19) Wada, Y.; Tajiri, M.; Yoshida, S. Hydrophilic affinity isolation and MALDI multiple-stage tandem mass spectrometry of glycopeptides for glycoproteomics. *Anal. Chem.* **2004**, *76*, 6560–6565.
- (20) Selman, M. H. J.; Hemayatkar, M.; Deelder, A. M.; Wührer, M. Cotton HILIC SPE microtips for microscale purification and enrichment of glycans and glycopeptides. *Anal. Chem.* **2011**, *83*, 2492–2499.
- (21) Wührer, M.; de Boer, A. R.; Deelder, A. M. Structural glycomics using hydrophilic interaction chromatography (HILIC) with mass spectrometry. *Mass Spectrom. Rev.* **2009**, *28*, 192–206.
- (22) Mysling, S.; Palmisano, G.; Højrup, P.; Thaysen-Andersen, M. Utilizing ion-pairing hydrophilic interaction chromatography solid phase extraction for efficient glycopeptide enrichment in glycoproteomics. *Anal. Chem.* **2010**, *82*, 5598–5609.
- (23) Omaetxebarria, M. J.; Häggglund, P.; Elortz, F.; Hooper, N. M.; Arizmendi, J. M.; Jensen, O. N. Isolation and characterization of glycosylphosphatidylinositol-anchored peptides by hydrophilic interaction chromatography and MALDI tandem mass spectrometry. *Anal. Chem.* **2006**, *78*, 3335–3341.
- (24) Gobom, J.; Nordhoff, E.; Mirgorodskaya, E.; Ekman, R.; Roepstorff, P. Sample purification and preparation technique based on nano-scale reversed-phase columns for the sensitive analysis of complex peptide mixtures by matrix-assisted laser desorption/ionization mass spectrometry. *J. Mass Spectrom.* **1999**, *34*, 105–116.
- (25) Osborne, T. B. The proteins of the wheat kernel. *The Carnegie Institution of Washington*; Carnegie Institute: Washington, DC, 1907; publication 84, pp 1–119.
- (26) Wührer, M.; Hokke, C. H.; Deelder, A. M. Glycopeptide analysis by matrix-assisted laser desorption/ionization tandem time-of-flight mass spectrometry reveals novel features of horseradish peroxidase glycosylation. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 1741–1748.
- (27) Prenner, C.; Mach, L.; Glossl, J.; Marz, L. The antigenicity of the carbohydrate moiety of an insect glycoprotein, honey-bee (*Apis mellifera*) venom phospholipase A2. The role of alpha 1,3-fucosylation of the asparagine-bound N-acetylglucosamine. *Biochem. J.* **1992**, *284*, 377–380.
- (28) Kaji, H.; Yamauchi, Y.; Takahashi, N.; Isobe, T. Mass spectrometric identification of N-linked glycopeptides using lectin-mediated affinity capture and glycosylation site-specific stable isotope tagging. *Nat. Protoc.* **2006**, *6*, 3019–3027.
- (29) Küster, B.; Mann, M. ¹⁸O-labeling of N-glycosylation sites to improve the identification of gel-separated glycoproteins using peptide mass mapping and database searching. *Anal. Chem.* **1999**, *71*, 1431–1440.
- (30) Hajkova, D.; Rao, K. C. S.; Miyagi, M. The pH dependency of the carboxyl oxygen exchange reaction catalyzed by lysyl endopeptidase and trypsin. *J. Proteome Res.* **2006**, *5*, 1667–1673.
- (31) Angel, P. M.; Lim, J. M.; Wells, L.; Bergmann, C.; Orlando, R. A potential pitfall in ¹⁸O-based N-linked glycosylation site mapping. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 674–682.

- (32) Lee, T. Y.; Huang, H. D.; Hung, J. H.; Huang, H. Y.; Yang, Y. S.; Wang, T. H. dbPTM: An information repository of protein post-translational modification. *Nucleic Acids Res.* **2006**, *34*, D622–627.
- (33) Lack, G. Clinical practice: Food allergy. *N. Engl. J. Med.* **2008**, *359*, 1252–1260.
- (34) O'Neil, C. E.; Lehrer, S. B. Occupational reactions to food allergens. In *The Food Allergy: Adverse Reactions to Foods and Food Additives*, 2nd ed.; Metcalfe, D. D., Sampson, H. A., Simon, R. A., Eds.; Blackwell Science: Cambridge, MA, 1997; pp 311–335.
- (35) Bardor, M.; Faveeuw, C.; Fitchette, A. C.; Gilbert, D.; Galas, L.; Trottein, F.; Faye, L.; Lerouge, P. Immunoreactivity in mammals of two typical plant glyco-epitopes, core alpha(1,3)-fucose and core xylose. *Glycobiology* **2003**, *13*, 427–434.
- (36) Westphal, S.; Kolarich, D.; Foetisch, K.; Lauer, I.; Altmann, F.; Conti, A.; Crespo, J. F.; Miranda, E. E.; Vieths, S.; Scheurer, S. Molecular characterization and allergenic activity of Lyc e 2 (β -fructofuranosidase), a glycosylated allergen of tomato. *Eur. J. Biochem.* **2003**, *270*, 1327–1337.
- (37) Hefferson, K. L. Recent advances in virus expression vector strategies for vaccine production in plants. *Sci. Rep.* **2012**, *1*, 1–5.
- (38) Palmisano, G.; Antonacci, D.; Larsen, M. R. Glycoproteomic profile in wine: A “sweet” molecular renaissance. *J. Proteome Res.* **2010**, *9*, 6148–6159.
- (39) Sander, I.; Rozynek, P.; Rihs, H. P.; van Kampen, V.; Chew, F. T.; Lee, W. S.; Kotschy-Lang, N.; Merget, R.; Bruning, T.; Raulf-Heimsoth, M. Multiple wheat flour allergens and cross-reactive carbohydrate determinants bind IgE in baker's asthma. *Allergy* **2011**, *66*, 1208–1215.
- (40) Kimoto, M.; Suzuki, M.; Komiyama, N.; Kunimoto, A.; Yamashita, H.; Hiemori, M.; Takahashi, K.; Tsuji, H. Isolation and molecular cloning of a major wheat allergen, Tri a Bd 27K. *Biosci., Biotechnol., Biochem.* **2009**, *73*, 85–92.
- (41) Pastorello, E. A.; Farioli, L.; Conti, A.; Pravettoni, V.; Bonomi, S.; Iametti, S.; Fortunato, D.; Scibilia, J.; Bindslev-Jensen, C.; Ballmer-Weber, B.; Robino, A. M.; Ortolani, C. Wheat IgE-mediated food allergy in European patients: α -Amylase inhibitors, lipid transfer proteins and low-molecular-weight glutenins. Allergenic molecules recognized by double-blind, placebo-controlled food challenge. *Int. Arch. Allergy Immunol.* **2007**, *144*, 10–22.
- (42) Sotkovsky, P.; Sklenar, J.; Halada, P.; Cinova, J.; Setinova, I.; Kainarova, A.; Golias, J.; Pavlaskova, K.; Honzova, S.; Tuckova, L. A new approach to the isolation and characterization of wheat flour allergens. *Clin. Exp. Allergy* **2011**, *41*, 1031–1043.
- (43) Larré, C.; Lupi, R.; Gombaud, G.; Brossard, C.; Branlard, G.; Moneret-Vautrin, D. A.; Rogniaux, H.; Denery-Papini, S. Assessment of allergenicity of diploid and hexaploid wheat genotypes: identification of allergens in the albumin/globulin fraction. *J. Proteomics* **2011**, *74*, 1279–1289.
- (44) Sanchez-Monge, R.; Garcia-Casado, G.; Lopez-Otin, C.; Armentia, A.; Salcedo, G. Wheat flour peroxidase is a prominent allergen associated with baker's asthma. *Clin. Exp. Allergy* **1997**, *27*, 1130–1137.
- (45) Matsuo, H.; Uemura, M.; Yorozuya, M.; Adachi, A.; Morita, E. Identification of IgE-reactive proteins in patients with wheat protein contact dermatitis. *Contact Dermatitis* **2010**, *63*, 23–30.
- (46) Lee, J.; Lim, K. T. SJSZ glycoprotein (38 kDa) modulates expression of IL-2, IL-12, and IFN- γ in cyclophosphamide-induced Balb/c. *Inflammation Res.* **2012**, *61*, 1319–1328.
- (47) Jin, C.; Bencúrová, M.; Borth, N.; Ferko, B.; Jensen-Jarolim, E.; Altmann, F.; Hantusch, B. Immunoglobulin G specifically binding plant N-glycans with high affinity could be generated in rabbits but not in mice. *Glycobiology* **2006**, *16*, 349–357.
- (48) Laštovičková, M.; Bobálová, J. MS based proteomic approaches for analysis of barley malt. *J. Cereal Sci.* **2012**, *56*, 519–530.
- (49) Perrocheau, L.; Rogniaux, H.; Boivin, P.; Marion, D. Probing heat-stable water-soluble proteins from barley to malt and beer. *Proteomics* **2005**, *5*, 2849–2858.
- (50) Konečná, H.; Müller, L.; Dosoudilová, H.; Potěšil, D.; Buršíková, J.; Sedo, O.; Márová, I.; Zdráhal, Z. Exploration of beer proteome using OFFGEL prefractionation in combination with two-dimensional gel electrophoresis with narrow pH range gradients. *J. Agric. Food Chem.* **2012**, *60*, 2418–2426.
- (51) Iimure, T.; Takoi, K.; Kaneko, T.; Kihara, M.; Hayashi, K.; Ito, K.; Sato, K.; Takeda, K. Novel prediction method of beer foam stability using protein Z, barley dimeric α -amylase inhibitor-1 (BDAI-1) and yeast thioredoxin. *J. Agric. Food Chem.* **2008**, *56*, 8664–8671.
- (52) Hao, J.; Li, Q.; Dong, J. J.; Yu, J. J.; Gu, G. X.; Fan, W.; Chen, J. Identification of the major proteins in beer foam by mass spectrometry following sodium dodecyl sulfate-polyacrylamide gel electrophoresis. *J. Am. Soc. Brew. Chem.* **2006**, *64*, 166–174.
- (53) Okada, Y.; Iimure, T.; Takoi, K.; Kaneko, T.; Kihara, M.; Hayashi, K.; Ito, K.; Sato, K.; Takeda, K. The influence of barley malt protein modification on beer foam stability and their relationship to the barley dimeric α -amylase inhibitor-I (BDAI-I) as a possible foam-promoting protein. *J. Agric. Food Chem.* **2008**, *56*, 1458–1464.
- (54) Iimure, T.; Nankaku, N.; Watanabe-Sugimoto, M.; Hirota, N.; Tiansu, Z.; Kihara, M.; Hayashi, K.; Ito, K.; Sato, K. Identification of novel haze-active beer proteins by proteome analysis. *J. Cereal Sci.* **2009**, *49*, 141–147.
- (55) Stanislava, G. A review: The role of barley seed pathogenesis-related proteins (PRs) in beer production. *J. Inst. Brew.* **2010**, *116*, 111–124.