High throughput quantitative glycomics and glycoform-focused proteomics of murine dermis and epidermis

Rie Uematsu[1,2], Jun-ichi Furukawa[1], Hiroaki Nakagawa[1], Yasuro Shinohara[1]*, Kisaburo Deguchi[1], Kenji Monde[1], Shin-Ichiro Nishimura[1]*

1) Division of Biological Sciences, Graduate School of Science

Frontier Research Center for Post-Genomic Science and Technology

Hokkaido University, Sapporo 001-0021 (Japan)

2) Basic Research Laboratory, Kanebo COSMETICS INC., Odawara 250-0002 (Japan)

* Correspondence: Shin-Ichiro Nishimura, Yasuro Shinohara

Division of Biological Sciences, Graduate School of Science

Frontier Research Center for Post-Genomic Science and Technology

Hokkaido University, Sapporo 001-0021 (Japan).

Phone and Fax: +81 11 706 9043,

Email: shin@glyco.sci.hokudai.ac.jp, yshinohara@glyco.sci.hokudai.ac.jp

Running Title: Glycomic approach for the focused proteomics

ABBREVIATIONS

aoWR, N^{α}-((aminooxy)acetyl)tryptophanylarginine methyl ester; CHCA, α-cyano-4-hydroxycinnamic acid; ConA, Concanavalin A; DHB, 2,5-dihydroxybenzoic acid; GU, glucose unit; LC, liquid chromatography; LG, lamellar granules; MALDI, matrix assisted laser desorption mass spectrometry; PA, 2-aminopyridine, PNGase F, peptide-*N*-glycosidase F; TOF, time-of-flight.

SUMMARY

Despite recent advances in our understanding of the significance of the protein glycosylation, the throughput of the protein glycosylation analysis is still too low to be applied for the exhaustive glycoproteomic analysis. Aiming to elucidate the *N*-glycosylation of murine epidermis and dermis glycoproteins, we employed here a novel approach for a focused proteomics. A gross *N*-glycan profiling (glycomics) of epidermis and dermis was first elucidated both qualitatively and quantitatively upon *N*-glycan derivatization with novel, stable isotope-coded derivatization reagents followed by MALDI-TOF(/TOF) analysis. This analysis revealed distinct features of the *N*-glycosylation profile of epidermis and dermis for the first time. A high abundance of high-mannose type oligosaccharides was found to be characteristic of murine epidermis glycoproteins. Based on this observation, we performed focused proteomics that carry high-mannose type glycans following direct tryptic digestion of protein mixtures and affinity enrichment. We identified 15 glycoproteins with 19 *N*-glycosylation sites that carry high-mannose type glycans by off-line LC-MALDI-TOF/TOF mass spectrometry. Moreover, the relative quantity of microheterogeneity of different glycoforms present at each *N*-glycan binding site was determined. Glycoproteins identified were often contained in lysosomes (e.g. cathepsin L, γ-glutamyl hydrolase), lamellar granules (e.g. glucosylceramidase, cathepsin D) and desmosomes (e.g. desmocollin 1, desmocollin 3, desmoglein). Lamellar granules are organelles found in the terminally differentiating cells of keratinizing epithelia, and desmosomes are intercellular junctions in vertebrate epithelial cells, respectively, thus indicating *N*-glycosylation of tissue-specific glycoproteins may contribute

to increase the relative proportion of high-mannose glycans. The striking roles of lysosomal enzymes in epidermis during lipid remodeling and desquamation may also reflect the observed high abundance of high-mannose glycans.

INTRODUCTION

The rapid progress in the sequence analysis of genomes of a variety of living organisms is accelerating the investigation of various related proteins involved in biological processes and disorders. Carbohydrate modifications can profoundly affect protein function. Their importance in disease is evident from a growing number of embryonic lethal phenotypes seen in knockout mice with defects in glycoconjugate assembly or processing, [1] and hence a large scale protein glycosylation analysis has become evidently important.

The accurate identification of protein glycosylation is challenging because of their complex structures, labile nature, and microheterogeneity. The presence of non-glycopeptides in a sample also limits the sensitivity for analysis of glycopeptides on mass spectrometry due to ion suppression. Although several new techniques enabling the large-scale identification of glycoproteins have recently been developed, [2-4] they can not afford to provide information about oligosaccharide moieties since the analysis is performed on peptides of which such moieties are enzymatically removed prior to the analysis. It is also well known that glycosylation is cell type specific, so a single glycoprotein can have a different spectrum of glycan structures when expressed in different cells. Recent progress in mass spectrometry (e.g. electron-capture induced dissociation using Fourier transform mass spectrometry, [5, 6] MALDI-LIFT-TOF/TOF [7]) demonstrated their ability to provide information both on peptide sequence and glycan structure for the analysis of glycopeptide. However, their throughput is not high enough to apply large scale protein glycosylation analyses. Therefore, unveiling the significance of protein glycosylation in an efficient manner requires further thought. One

solution is to develop a focused approach based on function and information content.

In this article, we describe a glycomic approach to rationalize the focusing process using murine dermis and epidermis as model studies. A gross *N*-glycan profiling of the tissue(s) of interest was first elucidated both qualitatively and quantitatively to grasp the characteristics of *N*-glycosylation profiles of the tissue(s). A rapid and sensitive quantitative *N*-glycosylation profiling technique based on stable isotope-coded, glycan-selective derivatization was newly developed for this purpose. Glycoproteins carrying unique oligosaccharides were then selectively analyzed following affinity enrichment. Considering glycoproteins are often involved in the adhesion of cells and their extracelluar matrices, glycoproteome of the epidermis, where living cells undergoes desquamation which comprises a major part of the epithelial barrier and are continuously being renewed, [8] may provide a new insight into the functional role of protein glycosylation.

Recently, comparative proteomic profiling of murine epidermis and subepidermal tissues, [9] and proteomic characterization of the plasma membrane of human epidermis [10] were successively reported. These studies successfully identified a number of epidermal proteins, but yielded no information on protein glycosylation. The structural elucidation of mammalian epidermal glycoconjugates has been studied mostly histochemically, using lectins [11] or monoclonal antibodies. [12] These studies revealed that cell surfaces of keratinocytes in the epidermis contain numerous glycoconjugates. However, in the histochemical approach, it is often difficult to differentiate whether the glycoconjugate of interest is a glycolipid, an *N*-glycosylated or an *O*-glycosylated glycoprotein. It has provided neither detailed structural

information of the oligosaccharides nor their carrier proteins.

In this study, we reveal that the *N*-glycosylation of epidermal glycoproteins is characterized in their markedly high levels of high-mannose type oligosaccharides by a novel, stable isotope assisted *N*-glycan profiling technique. Based on this observation, we identified glycoproteins that carry high-mannose type glycans by MALDI-TOF/TOF mass spectrometry following direct tryptic digestion of protein mixtures and affinity enrichment of the glycopeptides of interest. Our approach allowed rapid and sensitive quantitative glycomic profiling and it also provides not only the identification of glycoproteins carrying particular glycoforms, but also the determination of the *N*-glycosylation sites and the relative quantities of the microheterogeneous glycoforms present at each *N*-glycan binding site. Finally, the biological significance of the observed dominance of high-mannose type *N*-glycans will be discussed.

EXPERIMENTAL PROCEDURES

Materials. Male hairless mice (Hos/HR-1) were obtained from Nippon SLC (Shizuoka, Japan). They were fed a standard mouse diet and water ad libitum. 7-12 week old animals were used in this study. Trypsin, chymotrypsin and α-galactosidase were purchased from Sigma (St. Louis, MO). Alpha-mannosidase, β-*N*-acetylhexosaminidase, β-galactosidase from jack bean, Concanavalin A (ConA) agarose and pyridylaminated (PA)-isomalto-oligosaccharides were purchased from Seikagaku Co. (Tokyo, Japan). 2,5-Dihydroxybenzoic acid, human angiotensin II, bombesin and adrenocorticotropic hormone 18-39 were obtained from Bruker Daltonik (Bremen, Germany). Other materials were bought from the sources indicated: Sephadex G-15, Amersham Biosciences AB (Uppsala, Sweden); Bio-Gel P-4, Bio-Rad (Hercules, CA); sodium cyanoborohydride, Aldrich Chemical Co. (Milwaukee, WI); peptide-*N*-glycosidase F (PNGase F), Roche Molecular Biochemicals (Mannheim, Germany); pronase, Carbiochem (La Jolla, CA). Alpha-L-fucosidase from bovine kidney and other chemical reagents were obtained from Wako pure chemical industries, Ltd. (Osaka, Japan).

Synthesis of novel labeling reagents.

N^{α}-*((Bocaminooxy)acetyl)tryptophanylarginine Methyl Ester Hydrochloride.* To a solution of N^{α}- benzyloxycarbonyl)tryptophanylarginine methyl ester (ZWR-OCH_3) [20] (10 mg, 20 mmol) in MeOH (5mL) at room temperature was added 10% Pd/C (10 mg) and stirred under hydrogen atmosphere. After 2h, the reaction mixture was filtered through a membrane to

remove Pd/C and the filterate was concentrated to give WR-OCH_3. (positive-ion mode MALDI-TOF m/z 376). To a solution of (Boc aminooxy)acetic acid (172mg, 0.9 mmol) in dry THF (2mL) were added *N*-methylmorpholine (110 µl, 1.0 mmol) and isobutyl chloroformate (131 µl, 1.0 mmol) at -20 ºC. The mixture was heated to room temperature and stirred for 15 min, followed by the addition of a solution of the WR-OCH_3 (508 mg, 1.0 mmol) and $NaHCO_3$ (84 mg, 1.0 mmol) in H_2O (1 mL) at 0ºC. After stirring for 1h, solvents were removed under reduced pressure. The residue was chromatographed on a column of silica gel using 6:1 $CHCl_3$-MeOH to give the Boc-aoWR-OCH_3 (360 mg, 73 %).; ^{1}H NMR (500 MHz, CD_3OD) δ 7.60-6.98 (m, 5H, indole), 4.24 (s, 2H, $CH_2ONHBoc$), 3.62 (s, 3H, CH_3O); positive-ion mode MALDI-TOF m/z 547.

N^{α}-*((Aminooxy)acetyl)tryptophanylarginine Methyl Ester (aoWR(H)) Hydrochloride.* To a solution of $BocNHOCH_2COWR\text{-}OCH_3$ (18 mg, 33 µmol) in TFA (2 mL) at -20°C was stirred for 2h. The reaction mixture was evaporated and co-evaporated with toluene 3 times to give aoWR(H).; positive-ion mode MALDI-TOF m/z 447.

Arginine d_3*-Methyl Ester Dihydrochloride*. Thionyl choloride (1.4 mL, 20mmol) was added to CD_3OD (25 mL) at -10 ºC and the solution was stirred for 10 min. Arginine (1.1 g, 5 mmol) was added to a solution at room temperature and stirred for 2days. The reaction mixture was evaporated and co-evaporated with CD_3OD 3 times. The residue was dissolved in CD_3OD (3.5 mL), and diethyl ether (15 mL) was added. The crystalline product was collected, washed with diethyl ether, and dried to give the Arg-OCD_3 in quantitative yield; ^{1}H NMR (500 MHz, CD_3OD) δ4.10 (t, 1H, J = 6.5 Hz, α-CH), 3.3 (s, 2H, $CH_2ONHBoc$), 3.26 (t, 2H, J = 7.0 Hz,

CH_2NH).

Tissue preparation. Preparation of epidermis and dermis was performed according to the previously described procedure [13] with minor modifications. Briefly, full-thickness skin samples were taken from the dorsal area of mice. The epidermis was peeled from dermis by heat separation at 60 ˚C, 30 sec. Epidermis and dermis were minced and heated at 90 ˚C for 10 min in water, then defatted as described by Bligh and Dyer, [14] and lyophilized. For the glycomic analysis, sialic acid residues were hydrolyzed with HCl prior to defattening.

***N*-glycan release and chemical derivatization.** Each defatted and lyophilized tissue (equivalent to ~3 mg) was dissolved in 7 M guanidine hydrochloride, 0.5 M Tris-HCl (pH 8.5), 10 mM EDTA, reduced with dithiothreitol, *S*-carbamoylmethylated and dialyzed against 10 mM ammonium bicarbonate. Following deglycosylation by PNGase F treatment, samples were deproteinated by precipitation with acetonitrile. The supernatant was evaporated to dryness and redissolved in 30 uL of H_2O. An aliquot (10 uL) of dermis and epidermis samples were mixed with 10 uL of aoWR (D) and aoWR (H) (0.4 mM in 40 mM acetate buffer, pH 4.0), respectively, and were heated at 90 ˚C for 1h to prepare aoWR derivatized oligosaccharides. A weakly acidic condition (e.g. pH 4.0) was found to be recommended for the derivatization, since the methyl ester moieties of aoWRs could be partly deesterified under the weakly basic condition. Alternatively, the same samples were derivatized with 2-aminopyridine (PA) and sodium cyanoborohydride according to the procedure previously

described. [15-17] After removal of unreacted PA by Sephadex G-15, the PA-oligosaccharides were further purified by collecting the elution from 7 to 12 min from amide-80 (4.6 x 250 mm, Tosoh, Tokyo, Japan) using HPLC.

Gross *N*-glycan profiling by MALDI-TOF. The each aoWR derivatized sample (1 uL) was directly diluted with DHB (10 mg/mL in 30% acetonitrile) 100-fold without any further purification, and an aliquot (1 uL) was deposited on the stainless steel target plate. The aoWR(H) derivatized epidermis sample and aoWR(D) derivatized sample were first analyzed separately to elucidate the relative quantities of the different oligosaccharides present in each tissue. Each sample was then mixed in equal quantities and was subjected to MALDI-TOF analysis for the relative quantitation of dermis and epidermis oligosaccharides. MALDI-TOF data were obtained using an Ultraflex time-of-flight mass spectrometer (Bruker Daltonik, Bremen, Germany) equipped with a LIFT-TOF/TOF facility controlled by the FlexControl 2.0 software package. All of the spectra were obtained using a reflectron mode with an acceleration voltage of 25 kV, a reflector voltage of 26.3 kV, and a pulsed ion extraction of 160 ns in the positive ion mode. These were the results of the signal averaging of 1,000 laser shots. Signal intensity of each mass was automatically calculated by FlexAnalysis 2.0. Estimation of *N*-linked oligosaccharide structures was obtained by input of peak masses into the GlycoMod Tool (http://www.expasy.ch/cgibin/glycomod_form.html).

Gross *N*-glycan profiling by two-D mapping technique.

The PA derivatized oligosaccharides were analyzed according to the previously described procedure. [15] Briefly, the mixture of PA-oligosaccharides were applied to an octadecylsilyl-silica (ODS, 6 x 150 mm, Shimadzu, Kyoto, Japan) HPLC column, and the elution times of the individual peaks were normalized with reference to the PA-derivatized isomalto-oligosaccharides of polymerization degree 4-20 and represented by GU (ODS). Then, individual fractions separated on the ODS column were applied to the amide-80 column. Similarly, the retention time of the individual peaks on the amide-80 column were represented by GU (amide). Thus, a given compound from these two columns provided a set of GU (ODS) and GU (amide) values, which corresponded to co-ordinates of the two-dimensional sugar map. By comparison with the co-ordinates of reference PA-oligosaccharides, the *N*-glycans from skin were identified. Identification was confirmed by co-chromatography with a candidate reference on the columns and sequential exoglycosidase digestion. Molar ratios of *N*-glycans were calculated from the individual peak areas.

Preparation of glycopeptides. Defatted and lyophilized epidermis (50mg) was S-carbamoylmethylated. The alkylated proteins were dialyzed against 10 mM NH_4HCO_3, and were digested with trypsin. The digested proteins were applied to a ConA-agarose column (0.7 x 13 cm) equilibrated with 150 mM NaCl/ 10 mM Tris-HCl buffer pH 7.5. After washing the column with 5 mM methyl α-glucopyranoside in 150 mM NaCl/ 10 mM Tris-HCl buffer pH 7.5, the ConA-bound glycopeptides were eluted with 0.1 M methyl α-mannopyranoside in 150 mM NaCl/ 10 mM Tris-HCl buffer pH 7.5. The bound fraction was then separated by

ODS column using HPLC by a linear gradient of acetonitrile (0 to 32 %) in 0.1 % formic acid. Chromatography was carried out at a flow rate of 1ml/min at room temperature and was monitored at 214 nm. The glycopeptides mixture were separated into 100 fractions and dried with a centrifugal vacuum concentrator. The fractionated glycopeptides were dissolved in 10uL of 30% acetonitrile. A portion (1uL) of each fraction was deglycosylated by PNGase F and dissolved in the matrix solution.

Glycopeptide identification by MALDI-TOF/TOF. Each fraction with or without PNGase F treatment was mixed with DHB (10mg/mL in 30% acetonitrile) or CHCA (saturated solution in 0.1% trifluoroacetic acid) and was applied on the MALDI target plate. MALDI-TOF(/TOF) data was obtained using an Ultraflex time-of-flight mass spectrometer. In MALDI-TOF, spectra were obtained using a reflectron mode with an acceleration voltage of 25 kV, a reflector voltage of 26.3 kV, and a pulsed ion extraction of 160 ns in the positive ion mode. For fragmentation ion analysis in the tandem time-of-flight (TOF/TOF) mode, precursors were accelerated to 8kV and selected in a timed ion gate. Fragment ions generated by laser-induced decomposition of the precursor were further accelerated by 19kV in the LIFT cell and their masses were analyzed after passing the ion reflector. Masses were automatically annotated by using FlexAnalysis 2.0. External calibration of MALDI-TOF mass was carried out using singly charged monoisotopic peaks and fragments of a mixture of human angiotensin II (*m/z* 1046.542), bombesin (*m/z* 1619.823), adrenocorticotropic hormone 18-39 (*m/z* 2465.199).

Protein identification by database search. Peak lists were generated from the MS/MS spectra using Bruker flexAnalysis and were processed by the MASCOT algorithm (Matrix Science Ltd.) to assign peptide on the mouse genome sequence database. Database (MSDB) was searched for tryptic peptides with up to 1 miscleavage. All cystein residues were treated as being carbamoylmethylated. Deamidation of asparagines caused by deglycosylation was considered. We first screened the candidate peptides with probability-based Mowse scores that exceeded their thresholds ($P < 0.05$) and with MS/MS signals for y- or b-ions >5. If the peptide did not contain the consensus tripeptide sequence for *N*-linked glycosylation (NXS/T/C), the data was eliminated regardless of the matching score.

RESULTS

The analytical protocol. The employed analytical protocol is comprised of two steps (Scheme 1). First, whole *N*-glycans present in the epidermis and dermis were liberated and gross *N*-glycan profiles were elucidated both qualitatively and quantitatively by MALDI-TOF utilizing novel, stable isotope-coded derivatization reagents. Based on the observed *N*-glycan profile, affinity reagent to enrich particular type of glycopeptides was chosen. We employed a shotgun proteomic approach rather than a classical gel-electrophoresis based approach since many of the glycoproteins were known to be membrane associated. For the identification of glycopeptides, off-line LC-MALDI-TOF/TOF was employed. LC separation prior to MALDI-TOF/TOF was introduced aiming to detect as many glycopeptides as possible by reducing the complexity and ion suppression effect. Off-line connection of LC and MALDI-TOF/TOF was employed to allow separated glycopeptides individually digested by PNGase F. By introducing this process, one can be assured of the occurrence of *N*-glycosylation, the accurate size of *N*-glycan. It also improves the peptide identification efficiency.

Gross *N*-glycan profiling of murine epidermis and dermis. Aiming to establish a rapid, sensitive and quantitative *N*-glycosylation profiling technique, novel stable isotope-coded glycan selective derivatization reagents were developed. Based on our previous findings that incorporation of guanidino and hydrophobic functionalities significantly enhance the ionization efficiency of oligosaccharide, [18] tryptophanyl-arginine was employed as a scaffold.

The *N*-terminus of the dipeptide was used for the incorporation of oxylamino functionality, while the *C*-terminus was methyl esterificated to generate a deuterated (d_3-methyl) analog. As indicated in Fig. 1, oligosaccharides preferentially react with aoWRs to give corresponding oxims. In this study, epidermis and dermis samples were labeled with aoWR (H) and aoWR (D), respectively. As shown in Fig. 2a and b, the MALDI-TOF spectra obtained from each sample differed significantly. Glycomod search revealed that the *m/z* values of most of the signals observed were found to be identical to those of oligosaccharides known to be present in mammal (Table 1). All the assigned signals in Table 1 were confirmed to be aoWR derivatives by comparing the *m/z* differences when the labeling reagents for epidermis and dermis samples were exchanged (data not shown). The MALDI-TOF/TOF analysis of aoWR(H) and aoWR(D) derivatized oligosaccharides gave very intense fragment signal at *m/z* 431 and 434, respectively, that corresponds to the N-O bond cleavage of the oxime. Thus, this diagnostic signal could be also useful to distinguish between aoWR derivatives and contaminant signals. The same samples were also analyzed by a two-dimensional sugar mapping technique, and we confirmed that the *N*-glycan profiles derived from two different techniques were almost the same both qualitatively and quantitatively, except that MALDI-TOF analysis could detect far more *N*-glycans than two-dimensional mapping technique as indicated in Table 1. Detailed structures of major oligosaccharides were determined by sequential glycosidase digestions and were shown in the same table.

Relative quantitation between dermal and epidermal oligosaccharide profiles were performed by mixing the aoWR(H) and aoWR(D) derivatized samples in equal quantities and

subsequent MALDI-TOF analysis (Fig. 2c). By incorporating mass differences of 3 Da, the oligosaccharides from the isotope-encoded sample were easily distinguished from the corresponding natural isotopic abundance oligosaccharides in the mass spectrometer, allowing a direct quantitative measure of the tissue specific *N*-glycosylation (Fig. 2 d-f). Based on the relative signal intensities between the aoWR(H)- and aoWR(D)-derivatized oligosaccharides, the relative quantities of glycomics present in dermis and epidermis were determined as shown in Fig. 3.

The glycomic analysis of murine dermis and epidermis revealed distinct features of the *N*-glycosylation dermal and epidermal profiles. A noteworthy feature is that the murine epidermal glycoproteins are characterized by a high abundance of high-mannose type oligosaccharides. On the contrary, in dermal glycoproteins, the large majority of *N*-glycans are the complex type. This appeared to be unique because in vertebrates high-mannose type glycans are generally present at low levels, though they can normally be found. [19]

In this study, glycomic analyses were performed after the removal of sialic acid residues (see Experimental section) since possible occurrence of the metastable fragmentation of sialic acid residue(s) during the flight time of the ions could affect on the quantitative analysis of oligosaccharides. This inconsistency will be solved for instance by methyl esterification of the sialic acid residue(s) to render sialylated oligosaccharides chemically equivalent to neutral oligosaccharides as reported. [20] Development of rapid chemical modification procedure to stabilize sialic acid residue is also currently on progress in our laboratory.

Identification of glycoproteins carrying high-mannose type oligosaccharides in epidermis. Driven by the observed unique *N*-glycan profile in epidermal tissues, identification of the proteins that carry high-mannose type oligosaccharides was performed. ConA was used as an affinity reagent to selectively recover the glycopeptides that carry high-mannose type oligosaccharides. [21] As shown in Fig. 4, the lectin chromatography greatly reduced the complexity of the peptide mixture. Following the enrichment of glycopeptides of interest in a crude tryptic digest of epidermis, the ConA-bound fraction was further separated into ~100 fractions by reversed-phase chromatography.

Simultaneous identification of peptide and *N*-glycan moieties were possible for some of the glycopeptides by MALDI-TOF/TOF mass spectrometry. An example is shown in Fig. 5. The most intense fragment signals of glycopeptides were a cleavage of the chitobiose core and a ^{O,2}X-ring cleavage of the innermost GlcNAc. By way of these signature signals, the signal for the deglycosylated peptide ion could be readily identified. Besides cleavages of glycan moieties, cleavages of peptide bonds were also observed. A database search based of the deglycosylated peptide ion mass, the fragment mass and intensity values lower than the molecular weight of deglycosylated peptide ion identified this peptide to be YYHGELSYLNVTR of cathepsin D with high probability-based Mowse scores ($P < 0.05$). The TOF/TOF fragmentation of the glycan moiety further indicated that this peptide was modified with $Man_5GlcNAc_2$. To the authors' knowledge, this is the first successful direct

glycopeptide identification based on raw TOF/TOF data taken from one glycopeptide derived from a complex biological mixture.

Despite the preferable nature of direct glycopeptide identification by MALDI-TOF/TOF, glycopeptide identification was performed after PNGase F digestion to improve the identification efficiency due to the lack of suitable software for the glycopeptide identification. An example is shown in Fig. 6. Mass spectrum obtained from HPLC fractions often contained plural glycopeptides' ions. The characteristic signal patterns separated by 162u (hexose) were often observed for those peptides carrying different high-mannose type glycoforms on the same peptide backbone (e.g. *m/z* 2211.36, 2534.35, 2696.40, and 2858.45 in Fig. 6A). Following the PNGase F digestion, all those signals disappeared. Instead, a new signal appeared (e.g. at *m/z* 1318.90 in Fig. 6B). This observation evidently revealed that the peptide was *N*-glycosylated with $Man_3GlcNAc_2$, $Man_5GlcNAc_2$, $Man_6GlcNAc_2$, and $Man_7GlcNAc_2$. The peptide was identified as TGEINITSIVDR of desmoglein by MALDI-TOF/TOF analysis as shown in Fig. 6C.

We observed that glycopeptides of different glycoforms on a same peptide tended to elute in close proximity on the reversed-phase chromatography analyses; peptides modified with larger *N*-glycans eluted earlier and those with smaller *N*-glycans eluted later. Therefore, the relative quantitation of the microheterogeneity of different glycoforms present at a particular *N*-glycan binding site could be determined by comparing the signal intensities upon mixing equivalent volumes from each successive fraction containing the same peptide

backbone. As shown in Fig. 7, the microheterogeneity of TGEINITSIVDR of desmoglein was clearly elucidated.

Based on the described procedure, we identified and characterized the major glycopeptides carrying high-mannose type oligosaccharides in the epidermis and listed them in Table 1. Out of 15 glycoproteins identified in this study, 4 were bona fide transmembrane proteins. The non-transmembrane proteins were mostly designated as either extracellular (2 proteins) or lysosomal (6 proteins), thus belonging to two cellular compartments known to be enriched for glycoproteins.

DISCUSSION

We described a glycomic approach to rationalize the focused glycoproteome analysis of the murine epidermis. Upon qualitative and quantitative gross *N*-glycan profiling analysis, we observed distinct features of the *N*-glycosylation profile between the two tissues for the first time. The newly developed stable isotope assisted derivatization reagents allowed a rapid and sensitive *N*-glycan profiling on MALDI-TOF in a quantitative manner of both intra- and inter-samples. Owing to its high ionization efficiency of aoWR on MALDI-TOF analysis, aoWR derivatives could be highly selectively detected with minimum purification steps. We have recently reported that polymers displaying oxylamino groups provide a useful platform for the "glycoblotting", a high throughput oligosaccharide purification procedure in a crude mixture. [22, 23] Derivatization with aoWR is therefore regarded as an analogue of glycoblotting where chemoselective ligation is done toward functional low-molecular-weight compounds instead of polymers.

Among many qualitative and quantitative differences in *N*-glycan profiling, we focused on the high-mannose type *N*-glycans carrying glycoproteins in this study. Following the affinity enrichment, they were identified by off-line LC-MALDI-TOF/TOF mass spectrometry, which simultaneously provided information about the sugar binding site and the relative quantitation of microheterogeneity of different glycoforms present at each *N*-glycan binding site. Although the number of identified glycoproteins is so far limited, they are most likely to represent the major high-mannose glycan carrying proteins because the glycopeptide identification was carried out for those signals whose intensity was strongest.

Among glycoproteins identified in this study, almost half belong to lysosomal hydrolases. This observation appeared to be reasonable considering that the major sorting mechanism of those hydrolases to lysosome is by the mannose-6-phosphate pathway. [24] In the epidermis, enzymes of lysosomes and lamellar granules are present also in the extracellular compartment and are responsible for the lipid remodeling required to generate the barrier lamellae as well as for the reactions that result in desquamation. [25] Lamellar granules (LG), which are considered to be lysosome-related organelles, [8] are organelles present in the cytoplasm of cells in the granular layer and account for about 10 % of volume of the granular cell cytosol. [26] LG likely originate from the Golgi apparatus and are currently thought to be elements of the tubulo-vesicular *trans* Golgi network [27, 28] while it is not known whether these cells use the same sorting system for LG as that used for lysosomes. LG is major source of stratum corneum lipid precursors and various hydrolytic enzymes such as lipases, proteases, acid phosphatases and glycosidases [27, 29, 30] and other proteins including corneodesmosin. [31] Among lysosomal glycoproteins identified in this study, glucosylceramidase and cathepsin D are reported to be present in LG so far. [29, 31] Although further investigation is required, the high abundance of high-mannose type oligosaccharides observed in epidermis may be attributable to the striking roles of lysosomal enzymes and/or the high abundance of lamellar guranule in epidermis.

The desmoglein along with the desmocollins comprise the desmosomal cadherins, the extracellular domains of which proteins make up the extracellular core domain of the desmosome. [32] These proteins are expressed in a tissue-specific manner and provide the sticky

adhesion of the desmosome required between adjacent cells. Desmoglein 1 and 3, and desmocollin 1 and 3 expressions are restricted to certain specialized epithelia such as epidermal, tongue, tonsil and oesophagus tissues. Given these tissue-specific glycoproteins are dominantly modified with high-mannose glycans, it may also prove beneficial to increase the relative proportion of high-mannose glycans in epidermis. Although desmogleins and desmocollins are reported to be glycoproteins containing mainly *N*-glycans, detailed information about oligosaccharide structures has been so far limited. [33-35] The current study also raises an interesting query: why desmosome requires high-mannose glycans particularly, while it is possible that any type of oligosaccharides could protect desmosomal proteins from protease degradation and may prevent premature desquamation as suggested by Walsh and Chapman. [36]

Other extracellular glycoproteins identified in this study are lymphocyte antigen 6 complex locus G6C protein and extracellular matrix protein-1 (ECM1). The *Ly6g6c* gene is predicted to encode members of the Ly-6 superfamily of proteins based on translations of the predicted gene sequences. [37] Murine Ly-6 antigens were originally identified as markers for hematopoietic cells. Although the precise functions of the Ly-6 antigens have not been determined, several lines of evidence have suggested that they are involved in cell signaling and cell adhesion. [38] ECM1 is a secreted glycoprotein first isolated from an osteogenic mouse cell line in concomitance with studies on bone matrix biology. [39] Subsequently, the human homologue has been found to regulate endochondral bone formation, and to stimulate proliferation of endothelial cells and induce angiogenesis. The *Ecm1* gene has two splice

variants. The long isoform is expressed in a number of tissues including liver, heart and lungs, while the short isoform expression is confined to skin and cartilage-containing tissues such as tail and front paw. ECM1 may be involved in the control of keratinocyte differentiation, but its precise role within the epidermis is not clear. [40] To authors' knowledge, structural elucidations of *N*-glycans of theses glycoproteins have been scarce. The evidence that extracellular glycoproteins are actually modified with high-mannose type oligosaccharides indicates that epidermis cells project high-mannose glycans on the cell surfaces and may be involved in molecular recognition events. For example, the trimeric extracellular domain of langerin, a cell surface receptor unique to Langerhans cells, is reported to bind mammalian high-mannose oligosaccharides though the intrinsic ligand is not known. [41]

The described protocol allows for determination of the relative quantities of the microheterogeneous glycoforms present at each *N*-glycan binding site. It should be noted that relatively abundant $Man_3GlcNAc_2$ (M3), $Man_3GlcNAc_2Fuc$ (M3F) and $Man_4GlcNAc_2$ (M4) were typically observed for lysosomal proteins. High mannose oligosaccharides are assembled in the endoplasmic reticulum and *cis*-Golgi, and contain between five and nine mannose residues. Therefore, the observed M3, M3F and M4 should be considered to be the degraded products. This can be explained by the presence of glycosidase (e.g. α-mannosidase) in lysosomes [42] and LG. [43] A peptide from desmoglein exhibited the exceptional presence of M3 and M4. The mechanism of the occurrence of such degradation on the plasma membrane surface may need to be further elucidated. Other glycopeptides from extracellular glycoproteins are commonly modified with high-mannose oligosaccharides with between five

and nine mannose residues in a distinct, microheterogeneitic manner. In this study, two functionally unknown proteins were also identified, Riken cDNA 1100001H23 [44] and Riken cDNA 2310020A21, [45] as high-mannose type oligosaccharides carrying proteins. The *N*-glycosylation microheterogeneity analysis revealed that the former is modified with $Man_5GlcNAc_2$, $Man_6GlcNAc_2$, $Man_7GlcNAc_2$ and $Man_8GlcNAc_2$ while the latter is predominantly modified with M3 and M3F. These observations may indicate that the latter belongs to lysosomal proteins while the former does not.

In summary, we proposed and demonstrated the feasibility of glycomic profiling in developing a focused approach for glycoproteomics. Protocols utilizing novel, stable isotope-coded derivatization reagents coupled with MALDI/TOF and off-line LC-MALDI-TOF/TOF were developed as key techniques for the glycomic and glycoproteomic analysis. This study demonstrates that the gross *N*-glycan profiling prior to glycoprotein identification makes the particular glycoforms-focused approach functionally meaningful and greatly accelerates the data mining process.

ACKNOWLEDGEMENTS

The authors would like to thank Mr. Joseph Justin Mulvey for correcting English. This work was partly supported by SENTAN, JST.

REFERENCES

1. Yamashita T., Wada, R., Sasaki, T., Deng, C., Bierfreund, U., Sandhoff, K., Proia, R.L. (1999) A vital role for glycosphingolipid synthesis during development and differentiation. *Proc. Natl. Acad. Sci. USA* **96,** 9142-9147

2. Xiong, L., Andrews, D., Regnier, F. (2003) Comparative proteomics of glycoproteins based on lectin selection and isotope coding. *J Proteome Res.* 2003, **2,** 618-625

3. Zhang, H., Li, X.J., Martin, D.B., Aebersold, R., (2003) Identification and quantification of *N*-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat Biotechnol.* **21,** 660-666

4. Kaji, H., Saito, H., Yamauchi, Y., Shinkawa, T., Taoka, M., Hirabayashi, J., Kasai, K., Takahashi, N., Isobe, T., (2003) Lectin affinity capture, isotope-coded tagging and mass spectrometry to identify *N*-linked glycoproteins. *Nat. Biotechnol.* **21,** 667-672

5. Kjeldsen, F., Haselmann, K.F., Budnik, B.A., Sorensen, E.S., Zubarev, R.A. (2003) Complete characterization of posttranslational modification sites in the bovine milk protein PP3 by tandem mass spectrometry with electron capture dissociation as the last stage. *Anal. Chem.* **75,** 2355-2361

6. Hakansson, K., Chalmers, M.J., Quinn, J.P., McFarland, M.A., Hendrickson, C.L., Marshall, A.G., (2003) Combined electron capture and infrared multiphoton dissociation for multistage MS/MS in a Fourier transform ion cyclotron resonance mass spectrometer. *Anal Chem.* **75,** 3256-3262

7. Kurogochi, M., Nishimura, S.-I., (2004) Structural characterization of *N*-glycopeptides by

matrix-dependent selective fragmentation of MALDI-TOF/TOF tandem mass spectrometry. *Anal Chem.* **76,** 6097-6101

8. Madison, K.C. (2003) Barrier function of the skin: "la raison d'etre" of the epidermis. *J. Invest. Dermatol.***121,** 231-241

9. Huang, C.M., Foster, K.W., DeSilva, T., Zhang, J., Shi, Z., Yusuf, N., Van Kampen, K.R., Elmets, C.A., Tang, D.C., (2003) Comparative proteomic profiling of murine skin. *J. Invest. Dermatol.* **121,** 51-64

10. Blonder, J., Terunuma, A., Conrads, T. P., Chan, K.C., Yee, C., Lucas, D.A., Schaefer, C.F., Yu, L.R., Issaq, H.J., Veenstra, T.D., Vogel, J.C. (2004) A proteomic characterization of the plasma membrane of human epidermis by high-throughput mass spectrometry. *J. Invest. Dermatol.* **123,** 691-699

11. Mehul, B., Corre, C., Capon, C., Bernard, D., Schmidt, R.., (2003) Carbohydrate expression and modification during keratinocyte differentiation in normal human and reconstructed epidermis. *Exp. Dermatol.* **12,** 537-545

12. Dabelsteen, E., Buschard, K., Hakomori, S., Young, W.W., (1984) Pattern of distribution of blood group antigens on human epidermal cells during maturation. *J. Invest. Dermatol.* **82,** 13-17

13. Proksch, E., Elias, P.M., Feingold, K.R. (1990) Regulation of 3-hydroxy-3-methylglutaryl-coenzyme A reductase activity in murine epidermis. Modulation of enzyme content and activation state by barrier requirements. *J Clin Invest.* **85,** 874-882

14. Bligh, E.G., Dyer, W.J. (1959) A rapid method of total lipid extraction and purification.

Can. J.Biochem. Physiol. **37,** 911-917

15. Nakagawa, H., Kawamura, Y., Kato, K., Shimada, I., Arata, Y., Takahashi, N. (1995) Identification of neutral and sialyl *N*-linked oligosaccharide structures from human serum glycoproteins using three kinds of high-performance liquid chromatography. *Anal. Biochem.* **226,** 130-138

16. Takahashi, N., Nakagawa, H., Fujikawa, K., Kawamura, Y., Tomiya, N. (1995) Three-dimensional elution mapping of pyridylaminated *N*-linked neutral and sialyl oligosaccharides. *Anal. Biochem.* **226,** 139-146

17. Yamamoto, S., Hase, S., Fukuda, S., Sano, O., Ikenaka, T. (1989) Structures of the sugar chains of interferon-gamma produced by human myelomonocyte cell line HBL-38. *J. Biochem.* **105,** 547-555

18. Shinohara, Y., Furukawa, J.-i., Niikura, K., Miura, N., Nishimura, S.I. (2004) Direct *N*-glycan profiling in the presence of tryptic peptides on MALDI-TOF by controlled ion enhancement and suppression upon glycan-selective derivatization. *Anal. Chem.* **76,** 6989-6997

19. Varki, A., Cummings, R., Esko, J., Freeze, H., Hart, G., Marth, J., editors. (1999) *Essentials of Glycobiology.* 1st ed. Cold Spring Harbor Laboratory Press, NY

20. Powell, A.K., Harvey, D.J. (1996) Stabilization of sialic acids in N-linked oligosaccharides and gangliosides for analysis by positive ion matrix-assisted laser desorption/ionization mass spectrometry. *Rapid Commun. Mass Spectrom.* **10,** 1027-32

21. Kobata, A., Endo, T. (1992) Immobilized lectin columns: useful tools for the fractionation

and structural analysis of oligosaccharides. *J Chromatogr.* **597,** 111-122

22. Nishimura, S.-I., Niikura, K., Kurogochi, M., Matsushita, T., Fumoto, M., Hinou, H., Kamitani, R., Nakagawa, H., Deguchi, K., Miura, N., Monde, K., Kondo, H. (2005) High-throughput protein glycomics: combined use of chemoselective glycoblotting and MALDI-TOF/TOF mass spectrometry. *Angew. Chem. Int. Ed.,* **44,** 91-96

23. Niikura, K, Kamitani, R, Kurogochi, M., Uematsu, R., Shinohara, Y., Nakagawa, H., Deguchi, K., Monde, K., Kondo, H., Nishimura, S.-I. (2005) Versatile glycoblotting nanoparticles for high-throughput protein glycomics. *Chemistry* **11,** 3825-3834

24. Kornfeld, S., Mellman, I. (1989) The biogenesis of lysosomes. *Annu. Rev. Cell Biol.* **5,** 483-525

25. Igarashi, S., Takizawa, T., Takizawa, T., Yasuda, Y., Uchiwa, H., Hayashi, S., Brysk, H., Robinson, J.M., Yamamoto, K., Brysk, M.M., Horikoshi,T. (2004) Cathepsin D, but not cathepsin E, degrades desmosomes during epidermal desquamation. *Br. J. Dermatol.* **151,** 355-361

26. Elias, P. M., Friend, D.S. (1975) The permeability barrier in mammalian epidermis. *J. Cell Biol.* **65,** 180-191

27. Madison, K.C., Sando, G.N., Howard, E.J., True, C.A., Gilbert, D., Swartzendruber, D.C., Wertz, P.W. (1998) Lamellar granule biogenesis: a role for ceramide glucosyltransferase, lysosomal enzyme transport, and the Golgi. *J Investig Dermatol Symp Proc.* **3,** 80-86

28. Elias, P.M., Cullander, C., Mauro, T., Rassner, U., Komuves, L., Brown, B.E., Menon, G.K. (1998) The secretory granular cell: the outermost granular cell as a specialized secretory cell.

J Investig Dermatol Symp Proc. **3,** 87-100

29. Grayson, S., Johnson-Winegar, A.G., Wintroub, B.U., Isseroff, R.R., Epstein, E.H.Jr., Elias, P.M. (1985) Lamellar body-enriched fractions from neonatal mice: preparative techniques and partial characterization. *J. Invest. Dermatol.* **85,** 289-294

30. Freinkel, R.K., Traczyk, T.N. (1985) Lipid composition and acid hydrolase content of lamellar granules of fetal rat epidermis. *J Invest Dermatol.* **85,** 295-298

31. Serre, G., Mils, V., Haftek, M., Vincent, C., Croute, F., Reano, A., Ouhayoun, J. P., Bettinger, S., Soleilhavoup, J. P. (1991) Identification of late differentiation antigens of human cornified epithelia, expressed in re-organized desmosomes and bound to cross-linked envelope. *J Invest Dermatol.* **97,** 1061-1072

32. Garrod, D.R., Merritt, A.J., Nie, Z. (2002) Desmosomal adhesion: structural basis, molecular mechanism and regulation. *Mol Membr Biol.* **19,** 81-94

33. Kapprell, H.P., Cowin, P., Franke, W.W., Ponstingl, H., Opferkuch, H.J. (1985) Biochemical characterization of desmosomal proteins isolated from bovine muzzle epidermis: amino acid and carbohydrate composition. *Eur J Cell Biol.* **36,** 217-229

34. Penn, E.J., Hobson, C., Rees, D.A., Magee, A.I. (1987) Structure and assembly of desmosome junctions: biosynthesis, processing, and transport of the major protein and glycoprotein components in cultured epithelial cells. *J Cell Biol.* **105,** 57-68

35. Ortiz-Urda, S., Elbe-Burger, A., Smolle, J., Marquart, Y., Chudnovsky, Y., Ridky, T.W., Bernstein, P., Wolff, K., Rappersberger, K. (2003) The plant lectin wheat germ agglutinin inhibits the binding of pemphigus foliaceus autoantibodies to desmoglein 1 in a majority of

patients and prevents pathomechanisms of pemphigus foliaceus in vitro and in vivo. *J Immunol.* **171,** 6244-6250

36. Walsh, A., Chapman, S.J. (1991) Sugars protect desmosome and corneosome glycoproteins from proteolysis. *Arch. Dermatol. Res.* **283,** 174–179

37. Ribas, G., Neville, M., Wixon, J.L., Cheng, J., Campbell, R.D. (1999) Genes encoding three new members of the leukocyte antigen 6 superfamily and a novel member of Ig superfamily, together with genes encoding the regulatory nuclear chloride ion channel protein (hRNCC) and an N omega-N omega-dimethylarginine dimethylaminohydrolase homologue, are found in a 30-kb segment of the MHC class III region. *J. Immunol.* **163,** 278–287

38. de Nooij-van Dalen, A.G., van Dongen, G.A., Smeets, S.J., Nieuwenhuis, E.J., Stigter-van Walsum, M., Snow, G.B., Brakenhoff, R.H. (2003) Characterization of the human Ly-6 antigens, the newly annotated member Ly-6K included, as molecular markers for head-and-neck squamous cell carcinoma. *Int J Cancer.* **103,** 768-774

39. Bhalerao, J., Tylzanowski, P., Filie, J. D., Kozak, C. A., Merregaert, J. (1995) Molecular cloning, characterization, and genetic mapping of the cDNA coding for a novel secretory protein of mouse. Demonstration of alternative splicing in skin and cartilage. *J. Biol. Chem.* **270,** 16385-16394

40. Chan, I. (2004) The role of extracellular matrix protein 1 in human skin. *Clin Exp Dermatol.* **29,** 52-56

41. Stambach, N.S., Taylor, M.E. (2003) Characterization of carbohydrate recognition by langerin, a C-type lectin of Langerhans cells. *Glycobiology* **13,** 401-410.

42. Merkle, R.K., Zhang, Y., Ruest, P.J., Lal, A., Liao, Y.F., Moremen, K.W. (1997) Cloning, expression, purification, and characterization of the murine lysosomal acid α-mannosidase. *Biochim. Biophys. Acta*. **1336,** 132-146

43. de Vries, A.C., Schram, A.W., Tager, J.M., Batenburg, J.J., van Golde, L.M. (1985) An improved procedure for the isolation of lamellar bodies from human lung. Lamellar bodies free of lysosomes contain a spectrum of lysosomal-type hydrolases. *Biochim. Biophys. Acta.*. **837,** 230-238

44. Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., Hayashizaki, Y. (2000) RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. *Genome Res.* **10,** 1757-1771

45. Kawai, J., et al. (2001) Functional annotation of a full-length mouse cDNA collection. *Nature* **409,** 685-690

FIGURE LEGENDS

Scheme 1. General scheme for glycoform-focused proteomics.

Figure 1. Structures of aoWRs and the chemistry of adduct formation with oligosaccharides.

Figure 2. MALDI-TOF spectra showing the *N*-glycan profiles of murine dermis and epidermis glycoproteins. *N*-glycans derived from dermis and epidermis were derivatized with aoWR(D) and aoWR(H), respectively, and analyzed either independently (a: epidermis, b: dermis) or after mixing them in equal quantities (c). Magnified views of the mass spectrum (c) for the region of signals designated as 2, 8 and 20 are shown in d, e and f, respectively. Black and red arrows indicate aoWR(H)- and aoWR(D)-derivatives, respectively. Structures of the ions with numbers are shown in Table 1.

Figure 3. Relative quantitation of *N*-glycans present in dermis and epidermis glycoproteins.

Figure 4. Elution profiles of epidermal peptides digested with trypsin on the ODS column, (a) before and (b) after enrichment by ConA affinity chromatography.

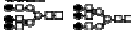
Figure 5. Tandem mass spectrum derived from TOF/TOF of the $[M+H]^+$ precursor, *m/z* 2831.6. The amino acids are represented in their single-letter code. The sequence identification of the peptide and database search showed that this glycopeptide was derived from cathepisn L, with the sequence of YYHGELSYLNVTR. The signal for the deglycosylated peptide is marked with asterisk. $^{0,2}X$, ring fragmentation of the innermost GlcNAc.

Figure 6. Identification of glycopeptide by MALDI-TOF/TOF following PNGase F

digestion of purified glycopeptides. MALDI-TOF spectra obtained for the purified glycopeptide with (A) and without (B) PNGaseF treatment. (C) Tandem mass spectrum derived from TOF/TOF of the $[M+H]^+$ precursor, *m/z* 1318.9. The sequence identification of the peptide and database search showed that this glycopeptide was derived from desmoglein, with the sequence of TGEINITSIVDR.

Figure 7. Determination of *N*-glycosylation microheterogeneity present at TGEINITSIVDR of desmoglein, (A) MALDI-TOF spectrum upon mixing the same volume from each successive HPLC fraction that contains same peptide backbone but with different glycoforms. (B) Relative quantitation of each glycoforms present at TGEINITSIVDR of desmoglein. Values are shown as mean ± SD (N=3). M3, $Man_3GlcNAc_2$, M3F, $Man_3GlcNAc_2Fuc$, M4, $Man_4GlcNAc_2$, M5, $Man_5GlcNAc_2$, M6, $Man_6GlcNAc_2$, M7, $Man_7GlcNAc_2$, M8, $Man_8GlcNAc_2$, M9, $Man_9GlcNAc_2$.

Table 1. Observed Signals of aoWR-Labeled Oligosaccharides Released from Dermis and Epidermis Glycoproteins

Signal no.	dermis m/z			epidermis m/z			composition/structure(s)	
	theor		obsd	theor		obsd		
1	1181.51		1181.44	1178.49		1178.35	(Hex)2 (HexNAc)2	
2	1327.57		1327.57	1324.55		1324.47	(Hex)2 (HexNAc)2 (dHex)1	
3				1340.55		1340.47	(Hex)3 (HexNAc)2	
4	1489.62		1489.67	1486.60		1486.59	(Hex)3 (HexNAc)2 (dHex)1	
5	1505.62		1504.72	1502.60	*	1502.60	(Hex)4 (HexNAc)2	
6	1546.65		1546.76	1543.63		1543.66	(HexNAc)1 + (Man)3(GlcNAc)2	
7				1648.66		1648.80	(Hex)4 (HexNAc)2 (dHex)1	
8	1667.67	*	1667.81	1664.65	*	1664.69	(Hex)2 + (Man)3(GlcNAc)2	
9				1705.68		1705.80	(Hex)1 (HexNAc)1 + (Man)3(GlcNAc)2	
10				1746.71		1746.78	(HexNAc)2 + (Man)3(GlcNAc)2	
11	1829.73	*	1829.90	1826.71	*	1826.77	(Hex)3 + (Man)3(GlcNAc)2	
12	1854.76		1854.95	1851.74		1851.79	(Hex)1 (HexNAc)1 (dHex)1 + (Man)3(GlcNAc)2	
13	1870.75	*	1870.95				(Hex)2 (HexNAc)1 + (Man)3(GlcNAc)2	
14	1895.78	*	1895.95				(HexNAc)2 (dHex)1 + (Man)3(GlcNAc)2	
15	1911.78	*	1911.95	1908.76	*	1908.84	(Hex)1 (HexNAc)2 + (Man)3(GlcNAc)2	
16				1988.76	*	1988.86	(Hex)4 + (Man)3(GlcNAc)2	
17				2013.79		2013.93	(Hex)2 (HexNAc)1 (dHex)1 + (Man)3(GlcNAc)2	
18	2032.80		2033.00				(Hex)3 (HexNAc)1 + (Man)3(GlcNAc)2	
19	2057.84	*	2058.03	2054.82	*	2054.96	(Hex)1 (HexNAc)2 (dHex)1 + (Man)3(GlcNAc)2	
20	2073.83	*	2074.01	2070.81	*	2070.95	(Hex)2 (HexNAc)2 + (Man)3(GlcNAc)2	
21				2111.84		2111.93	(Hex)1 (HexNAc)3 + (Man)3(GlcNAc)2	
22				2150.81	*	2150.91	(Hex)5 + (Man)3(GlcNAc)2	
23	2203.89		2202.20				(Hex)1 (HexNAc)2 (dHex)2 + (Man)3(GlcNAc)2	
24	2219.89	*	2220.07	2216.87	*	2216.99	(Hex)2 (HexNAc)2 (dHex)1 + (Man)3(GlcNAc)2	
25				2312.86	*	2312.99	(Hex)6 + (Man)3(GlcNAc)2	
26	2365.95		2365.19	2362.93		2363.02	(Hex)2 (HexNAc)2 (dHex)2 + (Man)3(GlcNAc)2	
27	2381.94		2381.28	2378.92		2379.10	(Hex)3 (HexNAc)2 (dHex)1 + (Man)3(GlcNAc)2	
28				2394.92	*	2395.05	(Hex)4 (HexNAc)2 + (Man)3(GlcNAc)2	
29	2438.96	*	2439.18				(Hex)3 (HexNAc)3 + (Man)3(GlcNAc)2	
30				2476.97		2477.12	(Hex)2 (HexNAc)4 + (Man)3(GlcNAc)2	
31				2508.99		2509.20	(Hex)2 (HexNAc)2 (dHex)3 + (Man)3(GlcNAc)2	
32	2528.00		2527.28	2524.98		2525.11	(Hex)3 (HexNAc)2 (dHex)2 + (Man)3(GlcNAc)2	
33				2540.98	*	2541.10	(Hex)4 (HexNAc)2 (dHex)1 + (Man)3(GlcNAc)2	
34	2585.02	*	2585.24				(Hex)3 (HexNAc)3 (dHex)1 + (Man)3(GlcNAc)2	
35	2950.15		2950.33				(Hex)4 (HexNAc)4 (dHex)1 + (Man)3(GlcNAc)2	

Oligosaccharides indicated with asterisk were detected also by two-dimensional mapping technique.
Structures were determined by two-dimensional mapping technique combined with sequential exoglycosidase digestion. □: GlcNAc, ○: Man, ●: Gal, ▽: Fuc

Table 2 Summary of identified glycoproteins in this study

Localization / Protein name *(gene name)*	Accession No. Swiss-Prot/TrEMBL	Sequence	Relative abundance (Mean ± SD) M3	M3F	M4	M5	M6	M7	M8	M9
extracellular										
Lymphocyte antigen 6 complex locus G6C protein *(Ly6g6c)*	Q9Z1Q4	LGLNYNTTCCDK	ND	ND	ND	22.5±1.4	29.9±2.1	24.8±0.6	22.8±14	ND
Desmocollin 1 *(Dsc1)*	P55849	NNQYNISVVATDTAGR	ND	ND	ND	54.3±1.6	24.6±0.8	17.4±1.0	3.67±0.8	ND
Desmocollin 3 *(Dsc3)*	P55850	IWTINQVNDTAAR	ND	ND	ND	18.8±2.4	34.1±0.5	42.2±3.3	4.9±0.7	ND
Desmoglein *(Dsg1gamma or Dsg6 or Dsg1b or Dsg5)*	Q7TQ61 or Q7TSF0 or Q7TQ60 or Q7TSF1	TGEINITSIVDR	7.3±0.9	ND	1.8±0.2	49.5±1.3	16.9±1.2	22.5±0.7	2.0±0.6	ND
extracellular matrix protein 1 *(Ecm1)*	Q61508	NVALVAGDTGNATGLGEQGPTR	ND	ND	ND	80.8±2.4	11.0±0.3	8.2±2.1	ND	ND
		QIPGLIQNMTVR	ND	ND	ND	60.4±10.2	39.6±10.2	ND	ND	ND
lysosome										
γ-glutamyl hydrolase *(Ggh)*	Q9Z0L8	KLPLNFTEGAR	60.0±2.5	3.9±0.2	1.8±0.3	29.8±2.4	3.2±0.3	1.2±0.1	ND	ND
		LPLNFTEGAR	65.5±1.2	4.2±0.5	1.7±0.4	24.1±2.4	1.7±0.2	2.9±1.1	ND	ND
		SINGVLLPGGGANLTDSGYSR	69.5±1.8	15.5±0.5	1.5±0.3	12.4±0.9	1.1±0.2	ND	ND	ND
		HFPTELLDSLALENLTANFHK	51.2±4.2	ND	8.9±1.1	33.7±1.0	6.2±3.0	ND	ND	ND
Cathepsin D *(Ctsd)*	P18242	YYHGELSYLNVTR	17.2±0.4	18.6±0.7	9.7±0.3	46.2±0.9	7.5±0.4	0.9±0.1	ND	ND
Lysosomal protective protein *(Ppgb)*	P16675	LDPPCTNTTAPSNYLNNPYVR	36.4±2.6	ND	2.8±0.5	54.5±2.7	6.4±0.2	ND	ND	ND
		MYVTNDTEVAENNYEALKDFFR	20.0±1.9	ND	9.5±1.3	62.3±2.2	8.2±0.7	ND	ND	ND
Cathepsin L *(Ctsl)*	P06797	AEFAVANDTGFVDIPQQEK	19.1±0.3	6.2±0.6	7.7±1.0	65.2±0.1	1.9±0.5	ND	ND	ND
Lysosome-associated membrane glycoprotein 1 *(Lamp1)*	P11438	AFNISPNDTSSGSCGINLVTLK	8.5±0.8	5.4±1.4	8.9±0.5	51.8±1.8	17.2±0.5	8.2±1.5	ND	ND
Glucosylceramidase *(Gba)*	P17439	VYTYADTPNDFQLSNFSLPEEDTK	65.1±3.4	19.3±1.7	ND	15.7±4.5	ND	ND	ND	ND
Tripeptidyl-peptidase I *(Cln2)*	O89023	DVGSGTTNNSQACAQFLEQYFHNSDLTEFMR	ND	ND	ND	69.5±5.3	30.5±5.3	ND	ND	ND
ER										
Signal sequence receptor, beta *(Ssr2)*	Q91Z43	IAPASNVSHTVVLRPLK	ND	ND	ND	ND	ND	17.2±1.7	41.6±2.5	41.3±1.8
Unknown										
RIKEN cDNA 1100001H23 *(1100001H23Rik)*	Q8VCI0	NGDAYGYYNDSIK	ND	ND	ND	27.5±2.4	35.1±4.8	23.4±4.9	14.0±2.0	ND
RIKEN cDNA 2310020A21 *(2310020A21Rik)*	Q8C1N0	WLVTNSTEALISHTLER	64.8±2.5	24.1±1.6	ND	11.2±2.4	ND	ND	ND	ND

Scheme 1, Uematsu et al.

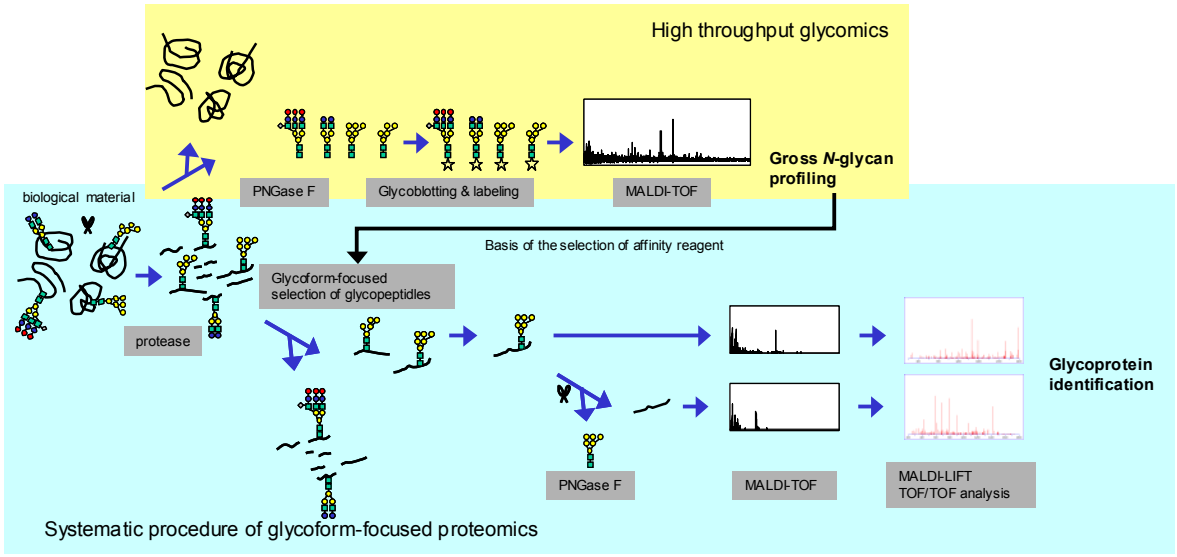


Figure 1, Uematsu et al.

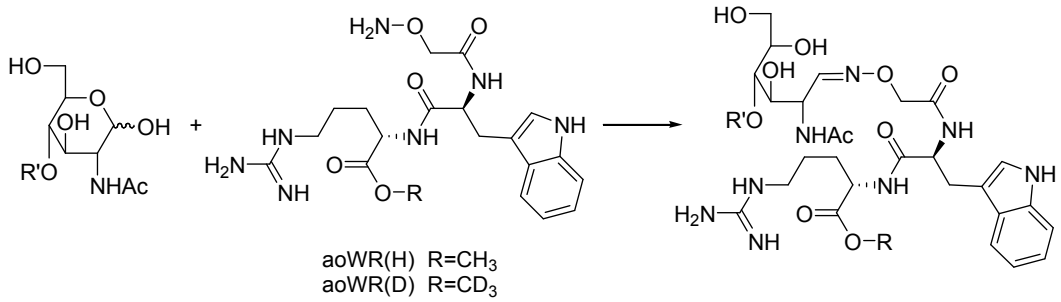


Figure 2, Uematsu et al.

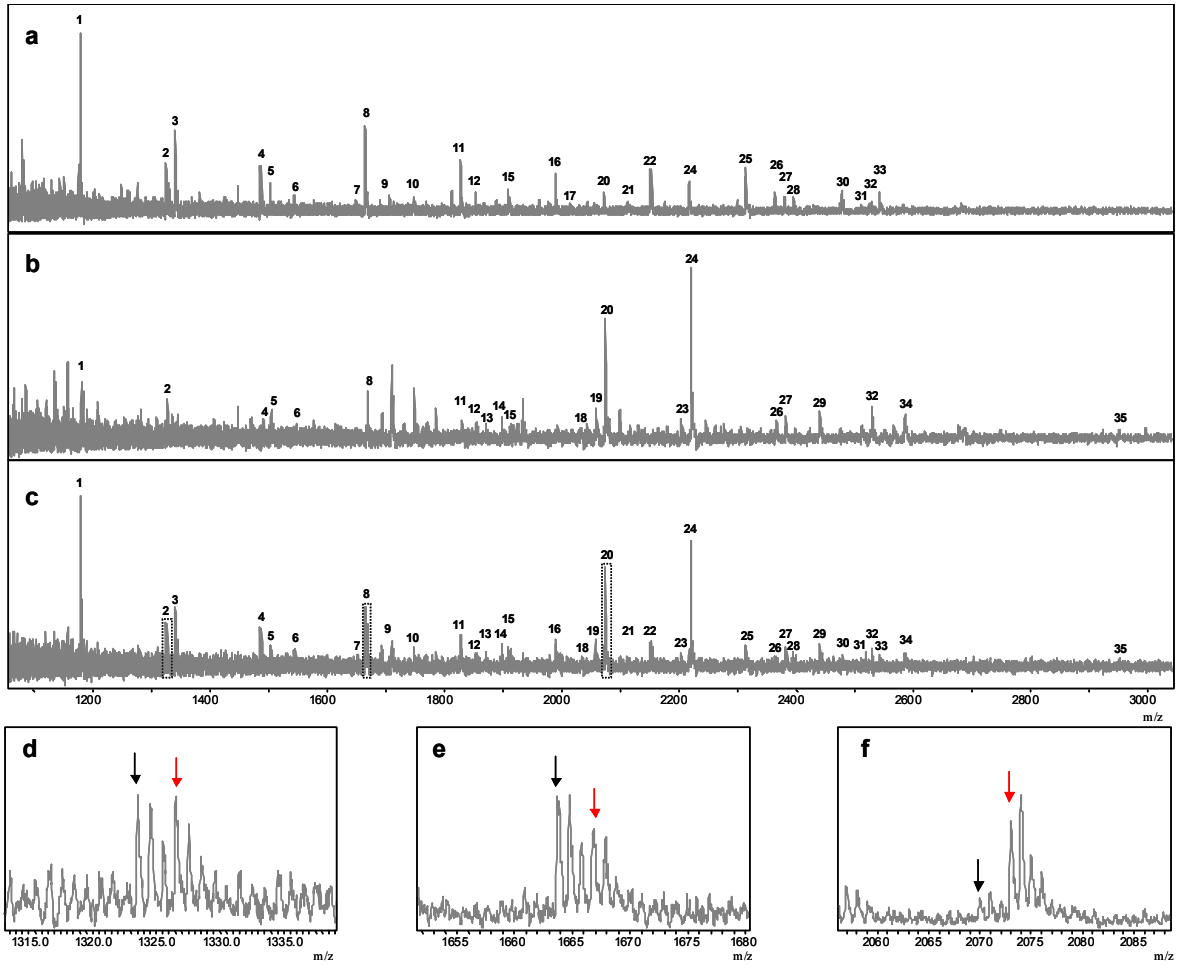


Figure 3, Uematsu et al.

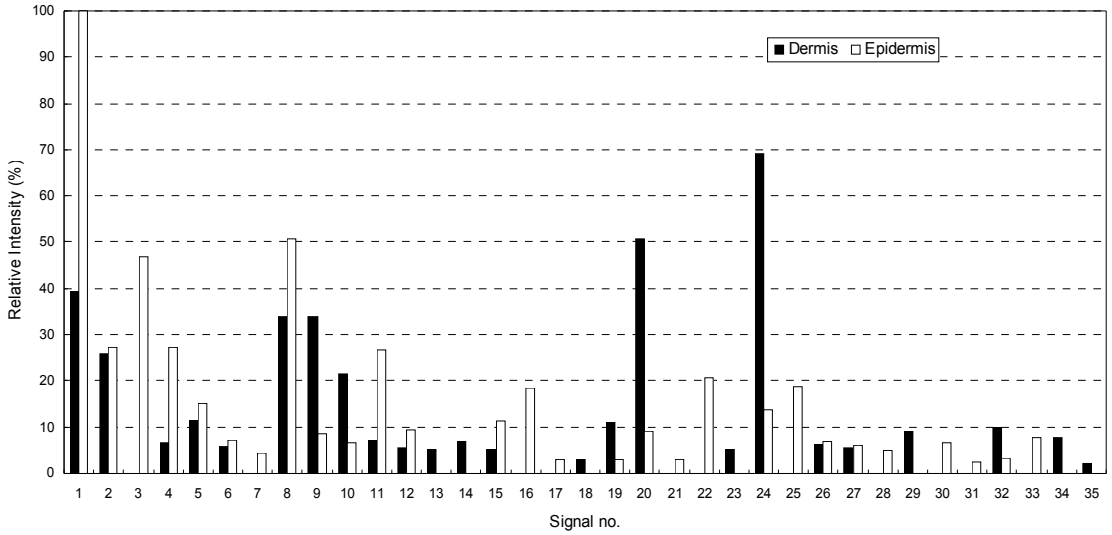


Figure 4, Uematsu et al.

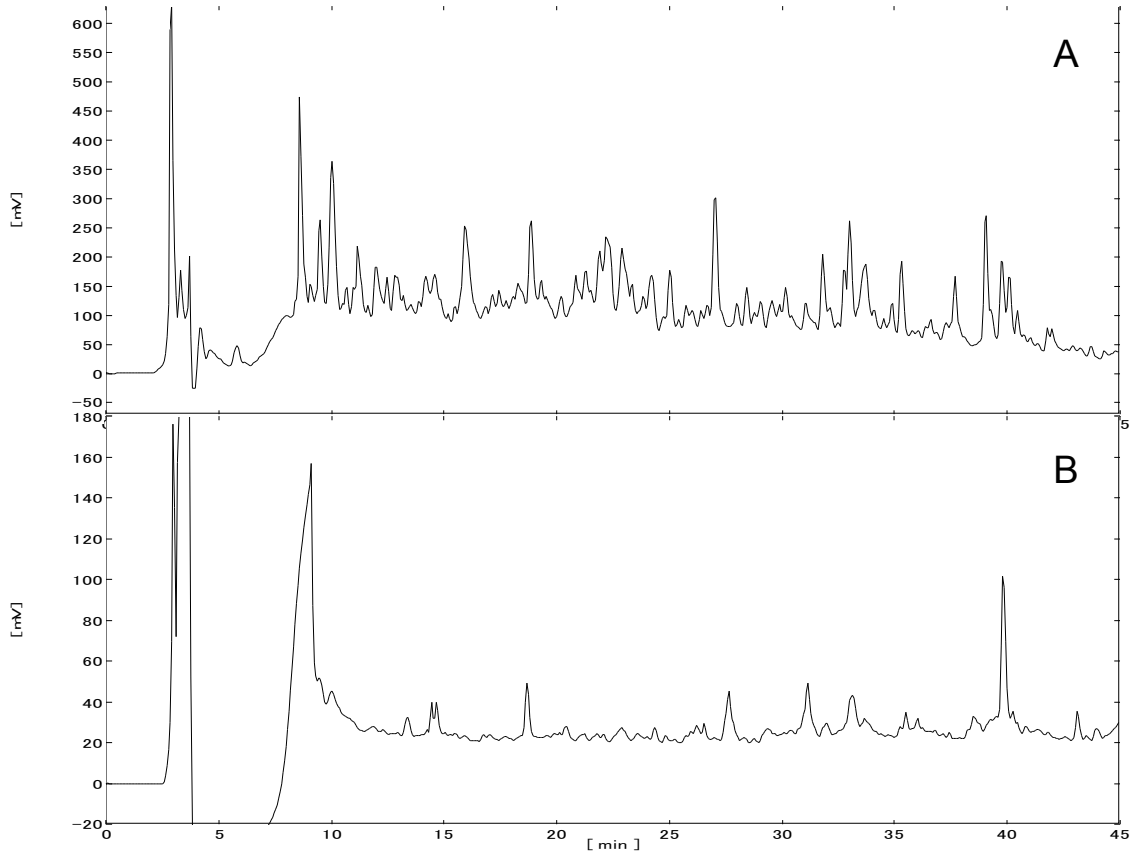


Figure 5, Uematsu et al.

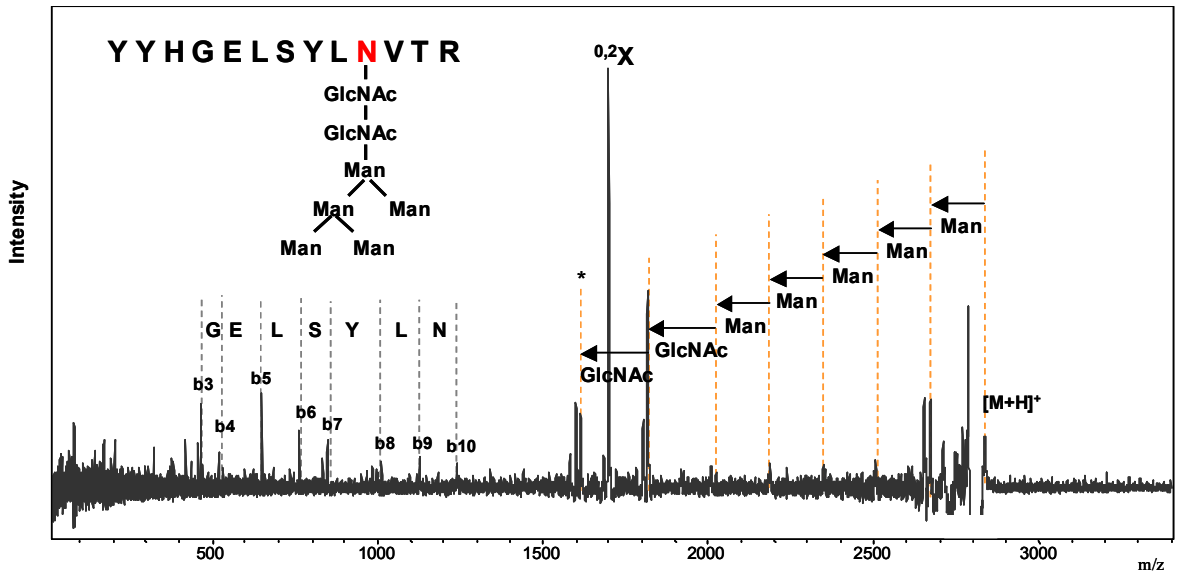


Figure 6, Uematsu et al.

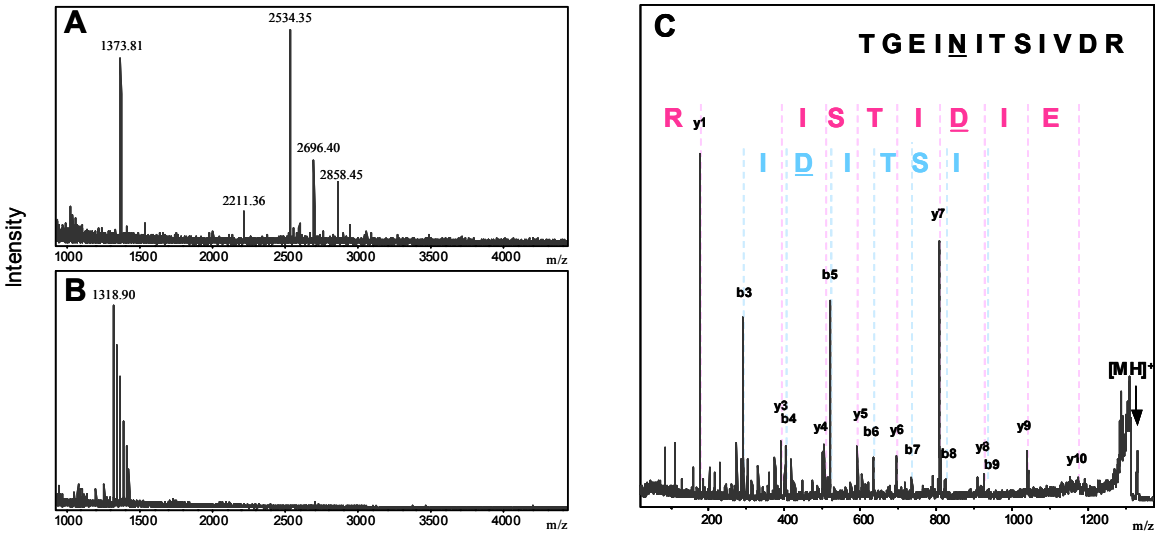


Figure 7, Uematsu et al.