Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

69451 Weinheim, Germany

Reverse glycoblotting allows rapid enrichment glycoproteomics of biopharmaceuticals and disease-related biomarkers

Masaki Kurogochi[1], Maho Amano[1], Masataka Fumoto[2], Akio Takimoto[2],

Hirosato Kondo[2], and Shin-Ichiro Nishimura[1,3]*

[1]Division of Advanced Life Science, Graduate School of Advanced Life Science,

Frontier Research Center for the Post-Genomic Science and Technology,

Hokkaido University, N21, W11, Sapporo 001-0021, Japan

[2]Shionogi & Co. Ltd., Osaka 553-0002, Japan

[3]Drug-Seeds Discovery Research Laboratory, National Institute of Advanced Industrial

Science and Technology (AIST), Sapporo, 062-8517 Japan

*Corresponding author: Dr. Shin-Ichiro Nishimura,

e-mail: shin@glyco.sci.hokudai.ac.jp

Experimental Section.

Materials. Bovine pancreas fibrinogen was purchased from Seikagaku Kogyo Co. Human fetal cord serum α-fetoprotein (AFP) was obtained from HyTest. Ltd. Recombinant human erythropoietin (rHuEPO, ESPO) was obtained from KIRIN BREWERY Co. Ltd. Sodium periodate ($NaIO_4$), trypsin (from porcine pancreas, mass spectrometry grade), α-chymotrypsin, and other reagents were obtained from Wako Pure Chemical. Peptide *N*-glycosidase F (PNGF) was obtained from Roche Applied Science. Male C57BL/KsJ *db/db* mice and age-matched *db/+* mice serum were purchased from Charles River Japan. Sephadex G-15 and 50 were purchased from Pharmacia Biotech Inc. MALDI matrices [2,5-dehydroxybenzoic acid (DHB) and α-cyano-4-hydroxycinnaminic acid (CHCA)] and MALDI peptide calibration standard mixture containing angiotensin II, bombesin, ACTH (18-39) and somatostatin were purchased from Bruker Daltonics GmbsH. Aminooxy functionalized polyacrylamide was synthesized by radical copolymerization using *N*-Boc protected acrylamide monomer derived from 6-aminohexanol and characterized according to the method previously reported (refs. [9] and [13]).

Selective and quantitative oxidation of glycoproteins. Fibrinogen (50 μg), AFP (50

μg), rHuEPO (30 μg), and mouse total serum (5 μl) were treated with 100 μl of 10 mM dithiothreitol (DTT) in 100 mM NH_4HCO_3 at 60ºC for 1 h, then the mixture was added and incubated with 100 μl of 50 mM iodoacetamide in 100 mM NH_4HCO_3 at 37ºC for 30 min. After precipitation by addition of acetone (1.8 ml), the alkylated glycoproteins were subjected to digestion with trypsin (for AFP) or a mixture of trypsin and α-chymotrypsin (for fibrinogen, rHuEPO, and mouse total serum) in 50 mM NH_4HCO_3 (the peptidase concentration was adjusted to be approximately 5 mg/ml and employed with a ratio of 1:50 against glycoproteins) at 37ºC for 16 h. After evaporating the solution, the residual peptidase digests were added with 50 μl of 3 mM $NaIO_4$ and reacted at 0 ºC for 15 min. The reaction was quenched by addition of the equivalent amount of 15 mM $Na_2S_2O_5$ (10 μl) to prevent the additional oxidation and the mixture was used directly for the next glycoblotting.

Enrichment of oxidized sialylglycopeptides by reverse glycoblotting. To the solution of oxidized glycopeptides (60 μl) was added 40 μl of Fischer-type polymer (final concentration of the polymer was adjusted to be 1 μg/μl), and the mixture was incubated at 37ºC for 2 h. The polymer solution was subjected to the purification by Sephadex G-50 column chromatography (0.6 x 10 cm, 3 ml) with water as eluent. The

isolated polymer fraction (approximately 1 ml) was mixed and treated with 300 μl of 10% trifluoroacetic acid (TFA) solution at 100ºC for 1 h to cleave the glycoside bond between sialic acid and galactose residues. Then, neutral asialo-glycopeptides were isolated by purification with Sephadex G-50 column chromatography (0.6 x 10 cm, 3 ml). To remove trace amount of TFA, crude products were subjected to co-evaporation with 50 mM NH_4HCO_3 using centrifugal evaporation system because a few pico-mole level of glycopeptides usually gives satisfactory MALDI-TOF mass spectra in the absence of TFA.

The modified protocol for for mice serum glycomics.

(1) Peptidase digests of glycoproteins prepared from 5 μl of mouse total serum are subjected to the treatment with protein *N*-glycanase F (PNGase F, 100 mU) at 37ºC for 3 h to release whole *N*-glycans from core peptides. (2) The mixtures of whole *N*-glycans and peptides are allowed to react with 1 mM sodium periodate at 0ºC for 15 min and subjected to the reverse glycoblotting with Fischer-type polymer reagent (1 μg/μl) at 37ºC for 2 h. Notably, the optimized condition employed for selective enrichment of aldehyde group of the oxidized sialic acid residues (37ºC for 2 h) does not work to react with any hemiacetal groups of common oligosaccharides. As

described in our previous reports (refs, [9], [19]-[21]), the glycoblotting for general sugar hemiacetals (free glycans) needs to be conducted at much higher temperature (80ºC) than that for aldehydes or reactive carbonyl groups (ref, [13]). To remove large excess of peptides and asialo-*N*-glycans, the polymer is purified by simple gel filtration with Sephadex G-50 column. (3) Isolated polymer is treated by 3% TFA at 100ºC for 1 h to release target oligosaccharides as asialo-*N*-glycans. Finally, purified asialo-*N*-glycans are analyzed and identified by MALDI-TOFMS.

MALDI-TOFMS and TOF/TOF analysis. DHB matrix solution was prepared with water at a concentration of 10 mg/ml. A few pico-mole of samples dissolved in water were applied on the target spot and mixed with 1μl of matrix solution, and dried at room temperature. MALDI-TOFMS were measured by using an Ultraflex TOF/TOF mass spectrometer equipped with a reflector, and controlled by the Flexcontrol 2.1 software package (Brucker Daltonics GmbsH, Bremen, Germany). In MALDI-TOFMS reflector mode, ions generated by a pulsed UV laser beam (nitrogen laser, λ =337 nm, 5 Hz) were accelerated to a kinetic energy of 23.5 kV. Metastable ions generated by laser-induced decomposition of the selected precursor ions were analyzed without any additional collision gas.

In MALDI-TOF/TOF mode (Ref. [14]), precursor ions were accelerated to 8 kV and selected in a timed ion gate. The fragments were further accelerated by 19 kV in the LIFT cell, and their masses were analyzed after the ion reflector passage. Masses were automatically annotated by using FlexAnalysis 2.1 software package. External calibration of MALDI mass spectra was carried out using singly charged monoisotopic peaks of a mixture of human angiotensin II (*m/z* 1046.542), bombesin (*m/z* 1619.823), ACTH (18-39) (*m/z* 2465.199), and somatostatin 28 (*m/z* 3147.472). Calibration of these mass spectra was performed automatically by utilizing a customized macro command of the FlexControl 2.1 software package. The macro command was used for the calibration of the monoisotopic singly charged peaks of the above mentioned peptides. TOF/TOF spectra were annotated with the BioTools 2.2 software package. All TOF/TOF spectra of mouse serum glycopeptides were basically searched by using the mouse International Protein Index (IPI) database (version 3.14 consisting of 160,914 protein entries; available online at http://www.ebi.ac.uk/IPI). The peptide fragment regions of the glycopeptides were used as searching fragments, which could be derived from the specific *N*-linked glycopeptide fragments observed as peptide fragments containing GlcNAc (203 *m/z*) and internal GlcNAc fragment moiety (83 *m/z*), and non sugar residues. The following dynamic modifications were also employed: (1)

carboxamidomethylation of cysteine residues and, (2) oxidation of methionine residues. When the peptides obtained by treating the original glycopeptides with PNGF were employed for searching by means of the MASCOT web search (http://www.matrixscience.com/mascot/), (3) the asparagine residues located in NXS/T motif were converted into aspartic acid residues. The MS/MS ion score was calculated using Bitools 2.2 (Bruker Daltonics).

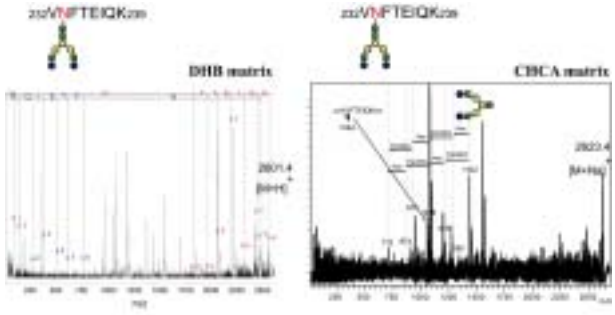
Conditions for the fractionation of enriched glycopeptides.

Total glycopeptides enriched from control and diabetes mice serum were subjected to the fractionation by means of gel filtration chromatography as follows.

Chromatography was conducted by using OH-pak802.5 column (Shodex, 7.0 x 300 mm) with 50 mM NH_4HCO_3 buffer as eluant at the flow rate of 0.2 ml/min. Monitoring was performed by a HITACHI L-6200 HPLC system equipped with a HITACHI L-7405 UV detector at 210 nm as shown in **Supplementary Fig. 7**. Each fraction (2 min/tube) was collected and concentrated by using centrifugal evaporation. Fractions 8-11 were subjected to MALDI-TOF mass analysis as indicated in **Supplementary Fig. 8**.

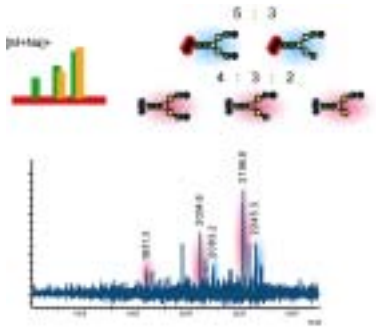
Kurogochi *et al*. Supplementary Fig. S-1

MALDI-TOF/TOF mass analysis of AFP by employing matrix dependent selective fragmentation (MDSF, ref. [14]). MDSF method allowed both glycan profiling and identification of glycosylation site concurrently.



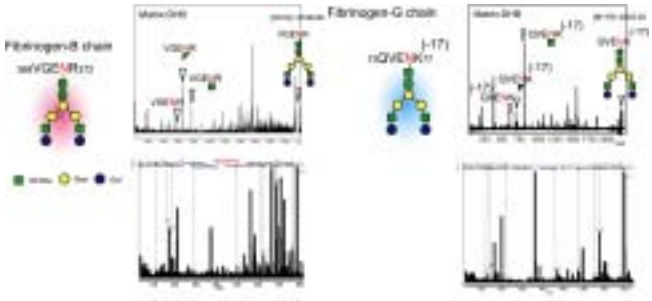
Kurogochi *et al*. Supplementary Fig. S-2

MALDI-TOF mass spectrum showing glycan microheterogeneity in two glycosylation sires of fibrinogen. Structural characterization of glycopeptides enriched from bovine pancreas fibrinogen yields 2 and 3 different glycoforms at ^{73}QVE<u>N</u>K^{77} and ^{368}VGE<u>N</u>R^{372} (Supporting Information **Fig. S-2, 3, 4, and 5**). Three ion peaks observed at *m/z* 2060.9, 1831.3, and 2034.8 were identified as peptides bearing a bi-antennary glycan chain involving an alternate terminal Man or GlcNAc residue, demonstrating that mono-sialylated glycans can also be enriched efficiently by the reverse glycoblotting technique.



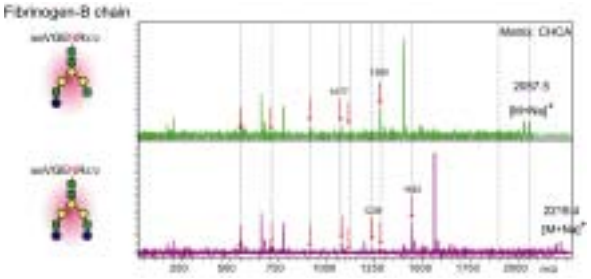
Kurogochi *et al*. Supplementary Fig. S-3

TOF/TOF mass analysis of individual glycopeptides for the identification both of glycosylation sites and glycoforms concurrently.



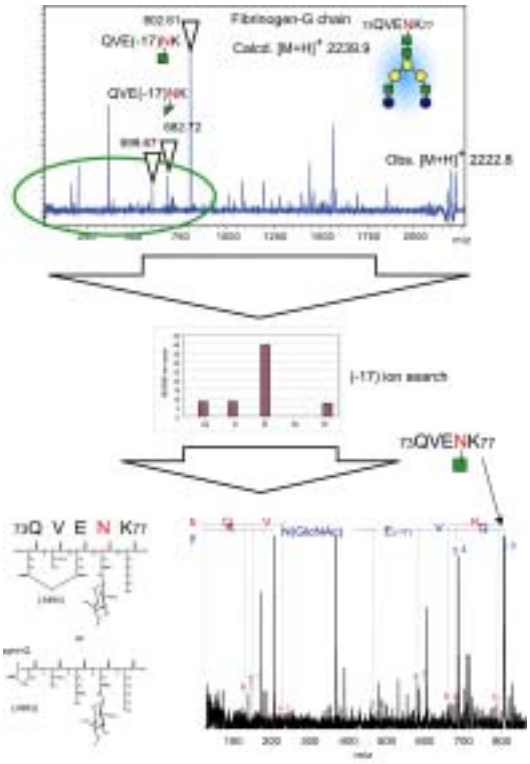
Kurogochi *et al*. Supplementary Fig. S-4

MDSF method allowed for glycan profiling of glycopeptides having the same peptide backbone. CHCA matrix-based TOF/TOF clearly showed the microheterogeneity of the terminal sugar residue, while DHB matrix was suited for proteomic approach.



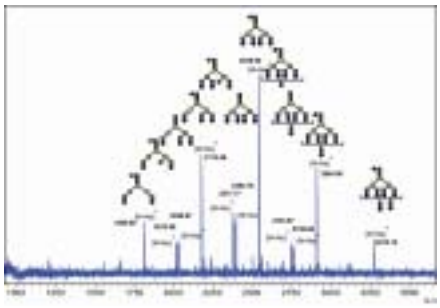
Kurogochi *et al*. Supplementary Fig. S-5

TOF/TOF mass analysis also revealed the existence of unique structure near the glycosylation site, QVEN*K, in fibrinogen G-chain.



Kurogochi *et al*. Supplementary Fig. S-6

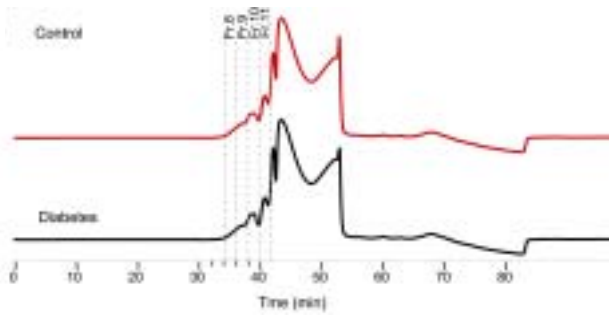
Glycan analysis of enriched rHuEPO glycopeptides were carried out on the basis of modified reverse glycoblotting method as described in the text. Precursor ion peaks were subjected to further TOF/TOF-based glycomics and identified.



We demonstrated that rHuEPO exhibits 8, 4, and 13 different glycoforms at ^{43}LLEAKEAE<u>N</u>I^{52}, ^{63}NE<u>N</u>I^{66}, and ^{108}LV<u>N</u>SSQPW^{115}, indicating that this method potentially makes it possible to identify the quality of biopharmaceuticals in terms of asialo-*N*-glycan microheterogeneity at the individual glycosylation site (**Fig. 2D**, see also Supporting Information **Fig. 6**).

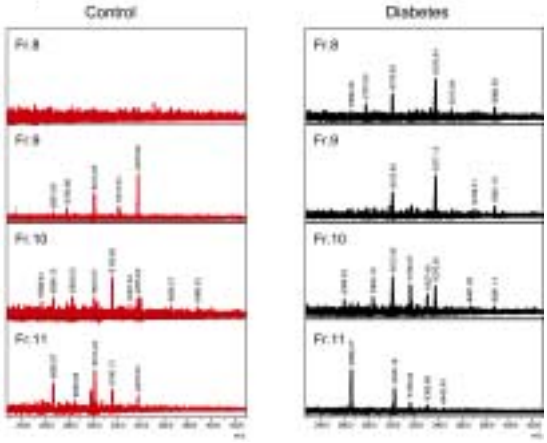
Kurogochi *et al*. Supplementary Fig. S-7

GFC patterns of mouse serum glycopeptides enriched by reverse glycoblotting.



Kurogochi *et al*. Supplementary Fig. S-8

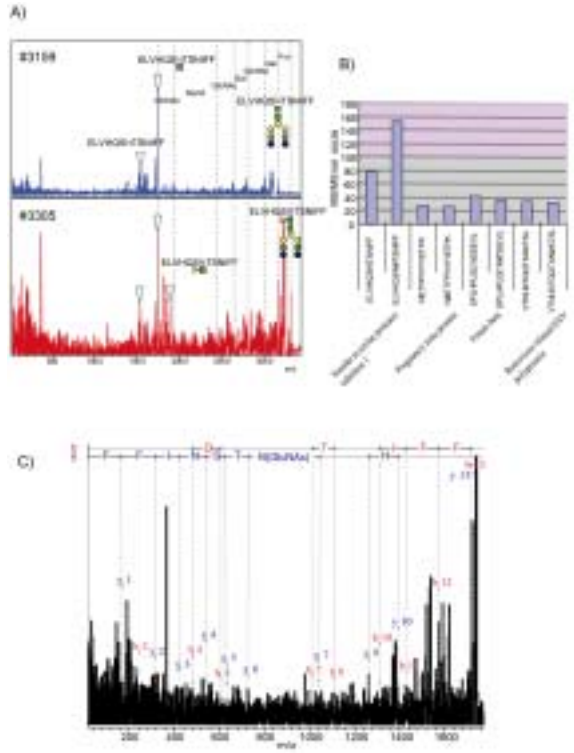
MALDI-TOF mass spectra of mouse serum glycopeptides after fractionation.



Kurogochi *et al*. Supplementary Fig. S-9

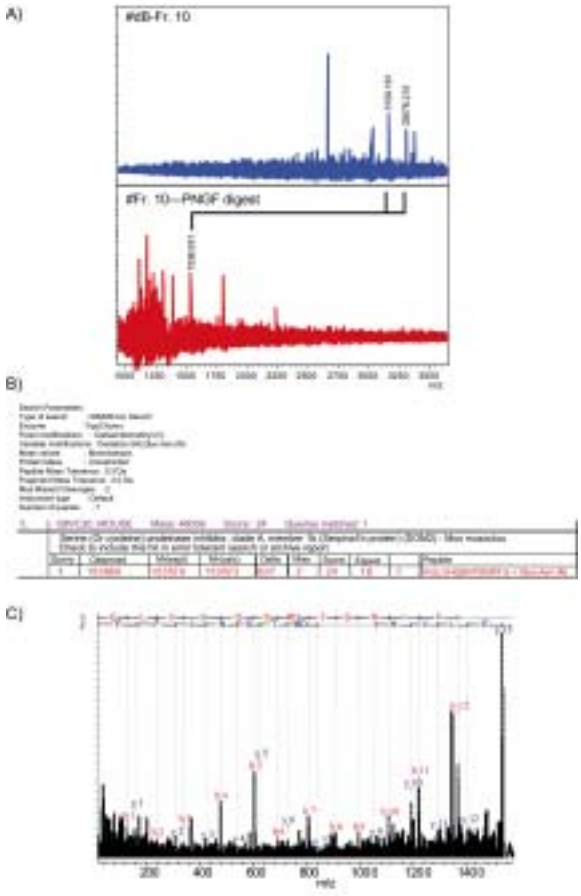
TOF/TOF analysis and Mascot searching of two definitive ion peaks found in fraction

10 indicated in Supplementary Fig. 8.



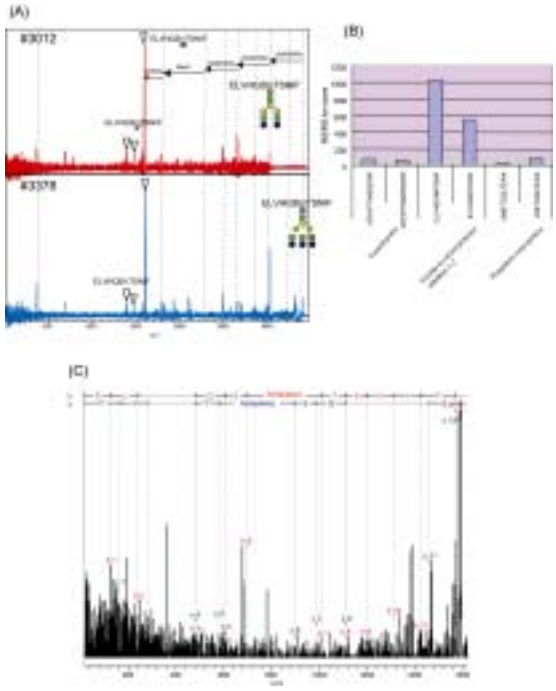
Kurogochi *et al*. Supplementary Fig. S-10

PNGF digest analysis of the parent glycopeptide observed at *m/z* 3159 and 3307 also revealed that the glycopeptides are arising from serine proteinase inhibitor 1.



Kurogochi *et al*. Supplementary Fig. S-11

It was also suggested that serine proteinase inhibitor 1 has two other glycan microheterogeneity at the same glycosylation site as indicated below. However, these glycoforms were detected equally both in control and diabetes mice shown in Supplementary Fig. 8.



Kurogochi *et al*. Supplementary Fig. S-12

Fractionation of the enriched glycopeptides may greatly facilitate further glycoproteomic analysis of relatively low abundance glycoproteins, even though the glycoforms identified here are not “focused glycoform” from the beginning.