A lectin-based isolation/ enrichment strategy for improved coverage of N-glycan analysis

Feng Guan^{a*}, Zengqi Tan^a, Xiang Li^b, Xingchen Pang^a, Yunlin Zhu^a, Dongliang Li^a, Ganglong Yang^a

^aThe Key Laboratory of Carbohydrate Chemistry & Biotechnology, Ministry of Education; School of Biotechnology, Jiangnan University, Wuxi, China.

^bWuxi Medical School, Jiangnan University, Wuxi, China.

* Corresponding author. Tel: +86-510-85918126. Fax: +86-510-85918126.
Email: fengguan@jiangnan.edu.cn

Abstract

Glycomics provides an increasingly useful research tool as the genomes and proteomes of more and more animal species are elucidated. In view of the general complexity and heterogeneity of glycans, improved depth-of-coverage and sensitivity are required for glycosylation analysis. In this study, we established the lectin-based isolation/enrichment strategy for total glycomic information. Specific lectins are added onto the filter to capture corresponding glycans prior to release of N-glycans by peptide N-glycosidase F (PNGase F). Non-bound glycans and bound glycans are released and analyzed by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS), respectively. Application of the strategy to chicken ovalbumin, normal mouse mammary epithelial cells (NMuMG), and human serum resulted in detection of 5, 6, and 11 additional N-glycan structures, respectively. The strategy facilitates identification of intact N-glycans in biological samples, and can be extended to detailed analysis of O-glycome or glycoproteome.

Keywords: N-glycan; lectin-based method; depth-of-coverage; time-of-flight mass spectrometry.

Glycosylation, one of the common types of post-translational protein modification, plays essential roles in many biological processes, including protein folding [1], protein clearance [2], and cell-cell interactions [3]. Aberrant glycosylation is often associated with inflammation [4], congenital disorders [5], and cancers [4]. Expression of specific glycans is a promising biomarker for stem cells [6] and cancer cells [7]. Glycomics, the comprehensive study of all glycans expressed in a particular cell, tissue, or organism, provides an increasingly useful research tool as the genomes and proteomes of more and more animal species are elucidated.

The categories of glycans include N-glycans, O-glycans, glycosphingolipids (GSLs), and glycosaminoglycans (GAGs). A variety of techniques for N-glycan analysis have been well established [7-10]. N-glycans can be easily released by the endoglycosidic enzyme PNGase F (N-Glycosidase F), and analyzed by high-resolution mass spectrometry (MS) [7, 11] and high-performance liquid chromatography (HPLC) [9]. During N-glycosylation progression, high mannose-type structures are the precursors of complex- and hybrid-type N-glycans. Mixtures of N-glycans typically include all three of these types. High mannose-type N-glycans, which range from $\text{Man}_5\text{GlcNAc}_2$ to $\text{Man}_9\text{GlcNAc}_2$, are the most abundant structures in a variety of cultured cell lines, and dominate the N-glycan profiles of many biological samples [6, 11].

Safety evaluation of therapeutic monoclonal antibodies (mAbs) requires careful detection of nonhuman N-glycans. N-glycan G0F ($\text{Man}_5\text{GlcNAc}_4\text{Fuc}$) comprises >70% of N-glycans in Bevacizumab and Adalimumab, two recombinant humanized mAbs. This highly abundant structure may hinder detection of nonhuman N-glycans that induce antibody-dependent cellular cytotoxicity (ADCC) [12].

High mannose-type structures can be recovered in the flow-through fraction by passage through a C18 cartridge, while the less abundant complex-type N-glycans are retained by the cartridge and analyzed by MS [13]. However, this method does not reveal the hybrid-type N-glycan structures in the N-glycan pool, and does not allow identification of additional complex-type N-glycans. In view of the general complexity and heterogeneity of glycans, improved depth-of-coverage and sensitivity are required for glycosylation analysis in cancer research, clinical measurements, and antibody-based therapeutics.

For analysis of glycans and glycoproteins, lectins were widely facilitated on the extraction of specific glycosylated proteins from total proteins. Multi-lectin affinity column containing Concanavalin A (ConA), Wheat germ agglutinin and Jacalin was developed to isolate N-linked and O-linked glycan and study variation of glycosylation [14-16]. Agarose-bound lectin of SNA, LCA and others were commercially developed to extract specific glycosylated glycoproteins [17-19]. Therefore, the strategy using specific lectins on enrichment of glycoproteins or glycopeptides could be applied to the glycan analysis with depth-of-coverage and high sensitivity.

We previously used a filter aided sample preparation (FASP)-based method [20, 21] to profile N-glycans in serum and cells [11, 22]. In these biological samples, high mannose-type or multi-antennary N-glycans typically predominate over low-abundant N-glycan structures which may carry biologically crucial glycoepitopes. We describe here a modified FASP-based approach for characterization of intact N-glycans with high sensitivity and depth-of-coverage (Fig. 1). In this method, specific lectins are added onto the filter to capture corresponding glycans prior to release of N-glycans by PNGase F. Non-bound glycans are collected by centrifugation, and bound glycans are released by eluting sugar and collected by centrifugation. Both non-bound and bound

glycans are then subjected to matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS).

The model glycoprotein used for this study was chicken ovalbumin (OVA), which contains predominantly high mannose-type N-glycans and is widely used in glycomics studies [23, 24]. To reduce the dominance of high mannose-type N-glycans, ConA, which binds to oligomannosidic, hybrid-type, and di-antennary N-glycans [25], was used to isolate and enrich N-glycans on OVA. Three groups of N-glycans (control, non-bound, and bound) were separated, enriched, and subjected to MALDI-TOF-MS. Twenty-five N-glycans were identified in the control group, consistently with previous reports [23, 24, 26-28], while 30 distinct N-glycan structures were identified in the combined bound+non-bound groups (Fig. 2; Table S1 [S= Supporting Information]). Each of the 25 N-glycans found in the control group was also seen in the bound+non-bound groups. Five unique N-glycan structures were detected in the bound+non-bound groups: m/z 1567.394, 1622.531, 1770.198, 1987.580, and 2231.457. The structure with m/z 1662.531 (hybrid-type; $\text{Gal}_1\text{GlcNAc}_3\text{Man}_5$) identified in the bound group had affinity for ConA. The m/z 1770.198 structure (fucosylated hybrid-type; $\text{Gal}_1\text{GlcNAc}_3\text{Man}_5\text{Fuc}_1$) was observed in the non-bound rather than the bound group, indicating that binding was hindered by fucosylation of GlcNAc in the antennae. The m/z 2231.422 structure (triantennary complex-type; $\text{Gal}_3\text{GlcNAc}_6\text{Man}_3$) was observed only in the non-bound group, suggesting that multi-antennary N-glycans are not retained on ConA. MALDI-TOF/TOF-MS/MS was performed to obtain detailed information regarding substitutions and branching patterns of monosaccharide constituents. MS/MS spectra of seven N-glycans (m/z 1581.361, 1647.514, 1663.478, 1742.606, 1809.561, 1975.608, 2288.689) are shown in Fig. S1.

Normal mouse mammary epithelial (NMuMG) cells expressed predominantly high mannose-type N-glycans [11]. Of total N-glycan structures, 70% correspond to high mannose-type $\text{Man}_{5-10}\text{GlcNAc}_2$, preventing detection of low-abundant N-glycans. We applied the ConA strategy as above to analyze N-glycans of NMuMG cells. The control group and bound+non-bound groups contained 25 and 29 distinct N-glycan structures, respectively (Fig. 3; Table S2). Following isolation and concentration of ConA, the relative intensity of high mannose-type N-glycans increased from 70% (control group) to 95% (bound group), and fell to 6% in the non-bound group. Twenty-three N-glycans were observed in both the control group and bound+non-bound groups. The bound+non-bound groups contained six unique N-glycan structures: m/z 1499.837, 1622.857, 1891.692, 1936.039, 2378.002, and 2488.098. Of these, three (m/z 1499.837, 1622.857, 1936.039) found only in the bound group were identified as biantennary hybrid-type. The other three (m/z 1891.692, 2378.002, 2488.098), found only in the non-bound group, were identified as multi-antennary, which cannot be retained on ConA. The noise intensity observed for the non-bound group was due to the relatively low-abundant complex-type N-glycans present in NMuMG cells (Fig. 3, Non-bound).

Human serum contains complex glycoproteins, which have been extensively studied in a search for glycan-associated disease biomarkers. MALDI-TOF-MS of serum N-glycans of the control group revealed high-abundant complex-type structures including terminal GlcNAc and internal lactosamine (Fig. 4). Four of these (m/z 1647.512, 1663.478, 1809.561, 2288.689), corresponding to lactosamine, accounted for 49.2% of total intensities (Table S3). We tested affinity for wheat germ agglutinin (WGA), which binds to internal lactosamine, terminal GlcNAc, and sialylated structures [29]. Twenty-nine and 38 N-glycans were found in the control group and

bound+non-bound groups, respectively. Of the 29 structures in the control group, 27 were also found in the bound+non-bound groups, giving a 93.1% coverage rate. The 11 structures unique to the bound+non-bound groups (m/z 1543.493, 1815.498, 1956.725, 1960.682, 2151.699, 2274.728, 2435.723, 2486.842, 2638.813, 2946.947, 3113.008) were lactosamine structures. Six structures were sialylated N-glycans (m/z 1960.682, 2435.723, 2486.842, 2638.813, 2946.947, 3113.008). Sialylated structures have been shown to be involved in regulation of many physiological processes, including molecular recognition [30] and survival of glycoproteins in the circulatory system [31]. The level of sialylation in sera of cancer patients is very different from that in healthy control subjects [32-34]. Our lectin-based strategy for serum analysis was able to reveal low-abundant sialylated N-glycan structures. The m/z 1956.725, 1960.682, 2486.842, 2638.813, and 3113.008 structures may be categorized as Lewis-type structures. Lewis-type antigens are often overexpressed in human carcinomas [35]. Lewis X (Le^x) antigen is a prognostic marker for several types of solid cancer, and a potential therapeutic target [36]. We observed multi-antennary structures (m/z 1543.493, 2151.699, 2274.728, 2435.723) only in the bound group. Altered expression of tri- and tetra-antennary structures has been reported in certain cancers [37, 38]. Our lectin-based strategy provides detailed information on serum N-glycans, with improved depth-of-coverage and sensitivity, and a molecular basis for elucidating the functions of low-abundant N-glycan structures.

The predominant N-glycan structures vary depending on the biological sample. Specific lectins can be selected for our strategy depending on the structures of the predominant N-glycans. For example, WGA and/or ConA can be used for N-glycan profiling when the predominant N-glycans are multi-antennary, sialylated, or high mannose-type. ConA and WGA are not glycoproteins (<http://www.vectorlabs.com/data/protocols/K4-K7.pdf>), which was further confirmed by MS (Fig. S2 and Fig. S3). In some cases, N-glycans of a biological sample may be contaminated by those released from other N-glycosylated lectins. A combination of two or three lectins would more reliably profile intact N-glycans in cases of predominant N-glycans having complex structures. For example, serial lectin affinity with *Sambucus nigra* agglutinin (SNA) and ConA has been used to isolate sialylated glycopeptides with complex biantennary N-linked, hybrid-type, and high mannose-type glycans [39]. Our lectin-based strategy can also be applied for analysis of other glycoconjugates, e.g., O-glycans. By combining this strategy with glycan derivatization, acetohydrazide could be introduced to modify sialic acid prior to release of glycans from glycoproteins [11, 40], and stable isotope tags such as [¹²C₆]- and [¹³C₆]-aniline could be used for quantitative glycan analysis [40].

In summary, the lectin-based strategy described here allows comprehensive detection of low-abundant N-glycan structures using a similar amount of proteins as in other methods, but with enhanced coverage and sensitivity. The strategy facilitates identification of intact N-glycans in biological samples, and can be extended to detailed analysis of O-glycome or glycoproteome.

1. Experimental

1.1. Cell culture

Mouse mammary epithelial cell line NMuMG, purchased from American Type Culture Collection (ATCC; Manassas, VA, USA), was cultured in DMEM containing 10% fetal bovine serum (HyClone; Logan, UT, USA), 10 µg/mL insulin (Sigma-Aldrich; St. Louis, MO, USA), 100

IU/mL penicillin, and 100 µg/mL streptomycin (Gibco; Carlsbad, CA, USA) at 37 °C in a 5% CO₂ atmosphere.

1.2. Release of N-glycans

Total proteins (2 mg of NMuMG and serum; 0.1 mg of ovalbumin) were concentrated using a size-exclusion spin ultrafiltration unit (Amicon Ultra-0.5 10 KD, Millipore; Billerica, MA, USA) at 12000 g. Proteins were denatured with 200 µL 8 M urea for 30 min at room temperature, reduced by adding an equal volume of 10 mM DTT (Sigma) at 56 °C for 45 min, and carboxyamidomethylated with 10 mM IAM (Sigma) at room temperature in darkness for 1 h. For the control group, N-glycans were directly released by PNGase F (New England BioLabs; Ipswich, MA, USA) in 200 µL 50 mM NH₄HCO₃. For the experimental group, each sample was washed with coupling buffer (0.1 mol/L Tris-HCl, 150 mmol/L NaCl, 1 mmol/L CaCl₂, 1 mmol/L MgCl₂, 1 mmol/L MnCl₂, pH 7.4) for 3 times, and 250 µg lectin in coupling buffer was introduced into the ultrafiltration unit and incubated for 3 h at 37 °C. The ultrafiltration unit was transferred to a new collection tube, and the sample was further digested with PNGase F overnight at 37 °C. Non-bound N-glycans were collected by centrifugation. Competitive inhibitors (0.5 mol/L α-methyl-mannose in coupling buffer for ConA; 0.1 mol/L N-acetyl-glucosamine in coupling buffer for WGA) were added to the ultrafiltration unit and incubated for 3 h at 37 °C. Bound N-glycans were collected. The non-bound and bound N-glycans were lyophilized.

1.3. Desalting of N-glycans

Sepharose 4B in a microtube was washed with methanol/ water (1:1, v/v) (MW) and 1-butanol/ methanol/ water (5:1:1, v/v) (BMW) under centrifugation. Glycans were dissolved in 500 µL BMW and added to the microtube. The mixture was gently shaken for 45 min and washed three times with BMW. N-glycans were eluted with MW, and the eluent was collected and lyophilized.

1.4. Mass spectrometry

N-glycans were characterized by MALDI-TOF/TOF-MS (UltrafleXtreme, Bruker Daltonics; Bremen, Germany). Lyophilized N-glycans were resuspended in 10 µL MW, and 1 µL of the mixture was spotted onto an MTP AnchorChip sample target and air-dried. One µL of 20 mg/mL 2,5-dihydroxybenzoic acid (DHB) in MW was spotted to recrystallize the glycans. Measurements were taken in positive-ion mode, and m/z data were analyzed and annotated using GlycoWorkbench software (<http://code.google.com/p/glycoworkbench/>). Relative intensity was calculated by dividing the average relative intensity of a given structure by the total average relative intensity of all N-glycans. Coefficient of variation (CV) percentages based on relative intensity values were used to estimate the stability of mass spectrometry.

Acknowledgments

This study was supported by the National Science Foundation for Young Scientists of China (No. 81201572), Jiangsu Province Recruiting Plan for High-level, Innovative and Entrepreneurial Talents, Jiangsu Province, “Six Summit Talent” Foundation (2013-SWYY-019), 111 Project (No. 111-2-06), and Fundamental Research Funds for the Central Universities (No. JUSRP51319B). The authors are grateful to Dr. S. Anderson for English editing of the manuscript.

References and notes

- [1] Parodi AJ. Annu Rev Biochem 2000;**69**:69-93.
- [2] Fukuda MN, Sasaki H, Lopez L, Fukuda M. Blood 1989;**73**:84-9.
- [3] Crocker PR, Feizi T. Curr Opin Struc Biol 1996;**6**:679-91.
- [4] Dube DH, Bertozzi CR. Nat Rev Drug Discov 2005;**4**:477-88.
- [5] Jaeken J, Matthijs G. Annu Rev Genom Hum G 2007;**8**:261-78.
- [6] Fujitani N, Furukawa J, Araki K, Fujioka T, Takegawa Y, Piao J, et al. Proc Natl Acad Sci U S A 2013;**110**:2105-10.
- [7] Yang G, Cui T, Wang Y, Sun S, Ma T, Wang T, et al. Proteomics 2013;**13**:1481-98.
- [8] Callewaert N, Schollen E, Vanhecke A, Jaeken J, Matthijs G, Contreras R. Glycobiology 2003;**13**:367-75.
- [9] Royle L, Campbell MP, Radcliffe CM, White DM, Harvey DJ, Abrahams JL, et al. Anal Biochem 2008;**376**:1-12.
- [10] Szabo Z, Guttman A, Rejtar T, Karger BL. Electrophoresis 2010;**31**:1389-95.
- [11] Tan Z, Lu W, Li X, Yang G, Guo J, Yu H, et al. J Proteome Res 2014;**13**:2783-95.
- [12] Kinoshita M, Kakehi K. Capillary electrophoresis of glycans: validation of antibody pharmaceuticals based on glycan profiles. In: Endo T, editor. *Glycoscience: Biology and Medicine*, Heidelberg: Springer; 2015. p. 111-8.
- [13] Lin CH, Kuo CW, Jarvis DL, Khoo KH. Proteomics 2014;**14**:87-92.
- [14] Yang Z, Hancock WS. J Chromatogr A 2004;**1053**:79-88.
- [15] Yang Z, Hancock WS, Chew TR, Bonilla L. Proteomics 2005;**5**:3353-66.
- [16] Yang Z, Harris LE, Palmer-Toy DE, Hancock WS. Clin Chem 2006;**52**:1897-905.
- [17] Nie S, Lo A, Wu J, Zhu J, Tan Z, Simeone DM, et al. J Proteome Res 2014;**13**:1873-84.
- [18] Tan Z, Yin H, Nie S, Lin Z, Zhu J, Ruffin MT, et al. J Proteome Res 2015;**14**:1968-78.
- [19] Wu J, Xie X, Nie S, Buckanovich RJ, Lubman DM. J Proteome Res 2013;**12**:3342-52.
- [20] Zielinska DF, Gnad F, Wisniewski JR, Mann M. Cell 2010;**141**:897-907.
- [21] Wisniewski JR, Zougman A, Nagaraj N, Mann M. Nat Methods 2009;**6**:359-62.
- [22] Yang G, Ma T, Li Z. Prog Biochem Biophys 2014;**41**:403-8.
- [23] Kim YG, Jeong HJ, Jang KS, Yang YH, Song YS, Chung J, et al. Anal Biochem 2009;**391**:151-3.
- [24] Jeong HJ, Kim YG, Yang YH, Kim BG. Anal Chem 2012;**84**:3453-60.
- [25] Wuhler M, Catalina MI, Deelder AM, Hokke CH. J Chromatogr B 2007;**849**:115-28.
- [26] Lattova E, Perreault H. J Chromatogr B 2003;**793**:167-79.
- [27] Harvey D, Wing D, Küster B, Wilson I. J Am Soc Mass Spectr 2000;**11**:564-71.
- [28] Saba JA, Shen X, Jamieson JC, Perreault H. J Mass Spectr 2001;**36**:563-74.
- [29] Porter A, Yue T, Heeringa L, Day S, Suh E, Haab BB. Glycobiology 2010;**20**:369-80.
- [30] Traving C, Schauer R. Cell Mol Life Sci CMLS 1998;**54**:1330-49.
- [31] Morell AG, Gregoriadis G, Scheinberg IH, Hickman J, Ashwell G. J Biol Chem 1971;**246**:1461-7.
- [32] Saldova R, Fan Y, Fitzpatrick JM, Watson RW, Rudd PM. Glycobiology 2011;**21**:195-205.
- [33] Kyselova Z, Mechref Y, Kang P, Goetz JA, Dobrolecki LE, Sledge GW, et al. Clin Chem 2008;**54**:1166-75.
- [34] Szabo Z, Guttman A, Karger BL. Anal Chem 2010;**82**:2588-93.

- [35] Dennis JW, Granovsky M, Warren CE. BBA-Gen Subjects 1999;**1473**:21-34.
- [36] Koh YW, Lee HJ, Ahn J-H, Lee JW, Gong G. Am J Clin Pathol 2013;**139**:746-53.
- [37] Goldman R, Ransom HW, Varghese RS, Goldman L, Bascug G, Loffredo CA, et al. Clin Cancer Res 2009;**15**:1808-13.
- [38] Alley Jr WR, Vasseur JA, Goetz JA, Svoboda M, Mann BF, Matei DE, et al. J Proteome Res 2012;**11**:2282-300.
- [39] Qiu R, Regnier FE. Anal Chem 2005;**77**:2802-9.
- [40] Yang G, Tan Z, Lu W, Guo J, Yu H, Yu J, et al. J Proteome Res 2015;**14**:639-53.

Figure legends

Fig. 1. Schematic workflow diagram of the lectin-based strategy used in this study.

Fig. 2. MALDI-TOF/TOF-MS spectra of N-glycans from ovalbumin. N-glycans from control group (released from glycoprotein without lectin), bound group (released from glycoprotein; bound with lectin) and non-bound group (released from glycoprotein; non-bound with lectin) were separated and desalted as described in M&M. MALDI-TOF-MS was performed in positive-ion mode. Representative N-glycan spectra from triplicate analyses are shown. Peaks (signal-to-noise ratio >5) were selected, and detailed structures were analyzed using GlycoWorkbench software. Proposed structure and m/z value are shown for each peak. N-glycans indicated by red boxes were newly identified by the lectin-based strategy.

Fig. 3. MALDI-TOF/TOF-MS spectra of N-glycans from NMuMG cells. N-glycans from control, bound, and non-bound groups were separated and desalted as described in M&M. Procedures and conventions as described for Fig. 2.

Fig. 4. MALDI-TOF/TOF-MS spectra of N-glycans from serum. N-glycans from control, bound, and non-bound groups were separated and desalted as described in M&M. Procedures and conventions as described for Fig. 2.

figure 1

figure 2

figure 3

figure 4

Graphical abstract