Carbohydrate Polymers 237 (2020) 116122

Contents lists available at ScienceDirect

Carbohydrate Polymers

journal homepage: www.elsevier.com/locate/carbpol

Multistage mass spectrometry with intelligent precursor selection for *N*-glycan branching pattern analysis

Chuncui Huang[a,1], Hui Wang[b,c,1], Dongbo Bu[b,c], Jinyu Zhou[a], Junchuan Dong[b,c], Jingwei Zhang[b,c], Huanyu Gao[a], Yaojun Wang[b], Wengang Chai[d,*], Shiwei Sun[b,c,*], Yan Li[a,c,*]

[a] Institute of Biophysics, Chinese Academy of Sciences, 15 Datun Road, Beijing 100101, China
[b] Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences, 6 Kexueyuan South Road, Beijing 100080, China
[c] University of Chinese Academy of Sciences, 19 Yuquan Road, Beijing 100049, China
[d] Glycosciences Laboratory, Faculty of Medicine, Imperial College London, London W12 0NN, United Kingdom

ARTICLE INFO

Keywords:
Intelligent precursor selection
N-gGlycan branching pattern
Multistage mass spectrometry
Distinctive fragments
Hypothesis and significance test

ABSTRACT

Biological functions of *N*-glycans are frequently related to their unique branching patterns. Multistage mass spectrometry (MS^n) has become the primary method for glycan structural analysis. However, selection of the best fragment as the precursor for the next round of product-ion scanning is important but difficult. We have previously proposed the concept and designed the approach of glycan intelligent precursor selection (GIPS) to guide MS^n experiments, but its use in *N*-glycans is not straightforward as some *N*-glycans are of high similarity in branching patterns. In the present work we introduced new elements to GIPS to improve its performance in *N*-glycan branching pattern analysis. These include a hypothesis and significance test, based on Bayes factor, and DP_{biased} as a new precursor selection strategy. The improved GIPS was successfully applied to identification of individual *N*-glycans, and incorporated into MALDI-MS *N*-glycan profiling for assignment of *N*-glycans obtained from glycoproteins and complex human serum.

1. Introduction

N-glycosylation is an important post translational modification of proteins. Within the immune system, *N*-glycans on the cell surface have a role in dictating the migration patterns of the cells and are involved in discrimination activities relevant to the pathophysiology of various autoimmune diseases. *N*-glycans can also direct trafficking of glycoproteins, and mediate cell-cell and cell-matrix interactions (Marth & Grewal, 2008; Paulson, Blixt, & Collins, 2006; Raman, Raguram, Venkataraman, Paulson, & Sasisekharan, 2005). These functions are closely related to the branching patterns of the *N*-glycan structures. For instance, an increase in *N*-glycans with bisecting *N*-acetylglucosamine (GlcNAc) occurs frequently in cancer cells (Anugraham et al., 2014; Kizuka & Taniguchi, 2016; Pompach et al., 2013), and large tetra-antennary branching structures are commonly formed in cancer tissues (Everest-Dass et al., 2016). Furthermore, dramatic increase of *N*-glycans lacking an outer arm galactose residue released from serum immunoglobulin G (IgG) was reported for patients with rheumatoid arthritis (RA) (Huhn, Selman, Ruhaak, Deelder, & Wuhrer, 2009; Pekelharing, Hepp, Kamerling, Gerwig, & Leijnse, 1988), and *N*-glycans with bisecting structures are important for differentiation of RA and osteoarthritis (Sun et al., 2019). Thus, *N*-glycan identification, especially detailed structure information including branching patterns, is essential for glycomic and biological function studies.

A number of analytical techniques have been used for glycans analysis, including high performance liquid chromatography (HPLC), capillary electrophoresis (CE), mass spectrometry (MS), and NMR (Aldredge, An, Tang, Waddell, & Lebrilla, 2012; Michael & Rizzi, 2015; Sarkar, Drouillard, Rivet, & Perez, 2015). Among these, MS is sensitive and can provide a wealth of structural information, particularly when tandem MS (MS/MS or MS^2) is used (Chai et al., 2006; Gao, Zhang, Liu, Feizi, & Chai, 2015) with or without the assistance of software and bioinformatic algorithm (Joshi et al., 2004; Ju et al., 2019; Lohmann & von der Lieth, 2004; Wang et al., 2019). Many different software tools (Hu, Khatrl, & Zaia, 2017; Rojas-Macias et al., 2019; Walsh, Zhao, Campbell, Taron, & Rudd, 2016) have been developed to assist interpretation of mass spectral fragmentation using different strategies, for example database search (Morimoto et al., 2015; Wang et al., 2020) and the so-called *de novo* method using sophisticated algorithms (Maass, Ranzinger, Geyer, von der Lieth, & Geyer, 2007). However, MS^2 is not

* Corresponding authors.
E-mail addresses: w.chai@imperial.ac.uk (W. Chai), dwsun@ict.ac.cn (S. Sun), yanli@ibp.ac.cn (Y. Li).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.carbpol.2020.116122
Received 19 October 2019; Received in revised form 10 February 2020; Accepted 3 March 2020
Available online 05 March 2020

always sufficient and multiple-stage MS (MS^n) may provide more insight into the detailed structure (Ashline, Singh, Hanneman, & Reinhold, 2005, 2007; Reinhold, Zhang, Hanneman, & Ashline, 2013; Zaia, 2008; Ashline, Zhang, & Reinhold, 2017). During product-ion scanning, multiple fragment ions are produced, and selection of a fragment as the precursor for subsequent scanning is important but it is not straightforward and the assistance of computer software is ideal, particularly for the non-expert laboratories.

We have previously designed such a computer-assisted, automated intelligent precursor-ion selection (GIPS) approach for MS^n to select the most 'structurally informative' peak as the precursor for the next round of scanning in order to produce a distinctive product-ion spectrum for identification of glycan branching pattern (Sun et al., 2018). GIPS assigns the structure of a glycan of interest by comparing its MS^n spectra with many possible glycans (candidate glycans) with identical masses listed in a well-established glycan database (e.g. CarbBank (Doubet & Albersheim, 1992)). When the molecular mass of the glycan of interest is identified from the primary mass spectrum, all glycans with identical molecular masses listed in the database are extracted and listed as possible glycans. Glycan identification is then achieved by comparison of its product-ion spectra at multiple stages with computer simulated MS^n spectra of all listed candidate glycans. The matched structure in the database is assigned as the glycan of interest. For GIPS, we proposed a new concept, the 'distinguishing power' (DP) of a fragment ion. DP is a calculated measure (Sun et al., 2018) of an ion's ability to produce a product-ion spectrum containing distinctive structural information to be used in its differentiation from other isomers. The higher the value, the more distinctive-structural information it can produce.

This approach has been successfully used to branching pattern analysis of various glycans and differentiation of isomeric structures, particularly for human milk oligosaccharides (Sun et al., 2018), and compared with the most intense peak selection (MIPS) strategy (Ashline et al., 2007) the currently used conventional method for precursor ion selection. However, as shown in the examples presented here, the branching patterns of *N*-glycans are of very high similarities and some isomeric *N*-glycan structures are difficult to be assigned even with MS^n and GIPS, as *N*-glycans do not normally have sufficient number of distinctive fragments at any stage of MS^n scanning for comparison with other candidates. We have now developed an improved strategy, the second generation of GIPS (GIPS-II), for *N*-glycan branching pattern analysis. Here we report the principle, the implementation and application of GIPS-II.

2. Materials and methods

2.1. Materials

N-Glycans A2G2S2, Hybrid-Octa, A3, A4, Man-6, A2, FA2, A2G2, FA2G2, Man-9, A2G2S1, FA2G2S1, FA2G2S2, A4G4, A3G3S3, Man-7D3 were from Ludger (Abingdon, England), and Man-7D1 was provided by Dr. Vladimir Piskarev (Nesmeyanov Institute of Organoelement Compounds, Russian Academy of Sciences, Moscow). A2G1 and FA2BG2 were purchased from Dextra Laboratories (Reading, England). Glycoproteins RNase B were from Sigma-Aldrich (St. Louis, MO), and mAb adalimumab is an over the counter drug. Human serum sample was collected from healthy humans, with the approval by the Biosafety and Ethics Committee of the Institute of Biophysics, Chinese Academy of Sciences, and pooled before analysis. Sodium chloride, disodium hydrogen phosphate, sodium phosphate monobasic sodium carbonate, hydrochloric acid, ammonium bicarbonate, sodium hydroxide, standard peptides, propanol, DL-dithiothreitol (DTT), iodoacetamide (IAA), methyl iodide, acetic acid and phosphate buffer were purchased from Sigma-Aldrich. Acetonitrile, ethanol and water were obtained from Avantor Performance Materials (Center Valley, PA). Trypsin and peptide N-glycosidase F (PNGase F) were purchased from Promega (Madison, WI). C18-Sep-Pak cartridge was obtained from Waters (Milford, MA). All the chemicals were of analytical grade or better, and used as received without further purification.

2.2. N-Glycan release

Human α-1-acid glycoprotein (AGP) was obtained from Sigma-Aldrich, and asialo-AGP was obtained by digestion of 200 μg AGP with 200 units of α-neuraminidase from *Arthrobacter ureafaciens* (Sigma-Aldrich) in 50 mM sodium acetate buffer (pH 5.5) at 37 °C overnight. RNase B, adalimumab and asialo-AGP were first reduced and carboxymethylated in the presence of DTT and IAA followed by dialysis in 50 mM Ambic buffer for 24 − 48 h at 4 °C. The glycoproteins were then digested by trypsin at 37 °C overnight. The resulting glycopeptides were applied to Sep-Pak C18 cartridge which was pre-conditioned successively with 5% acetic acid and 100 % propan-1-ol, and eluted stepwise with 20 % and 40 % propan-1-ol solution. The propan-1-ol fractions were digested with PNGase F at 37 °C for 20 − 24 h. Separation of the *N*-glycans from the digestion mixture was made by using the C18 Sep-Pak and the solvent system propan-1-ol/5% acetic acid.

2.3. Permethylation of glycans

Permethylation and purification were performed as previously reported (Dell et al., 1994). Briefly, methyl iodide was added to standard oligosaccharides or *N*-glycans released from glycoproteins in the presence of NaOH slurry, and the sample was then agitated on an automatic shaker at room temperature for 20 min. Chloroform was added and mixed thoroughly with the sample. The mixture was then allowed to settle into two layers. The upper aqueous layer was removed, and the chloroform layer was dried down under a gentle stream of nitrogen. Finally, the mixture was purified on a C18 Sep-Pak cartridge.

2.4. MALDI-MS

Permethylated glycan standards and *N*-glycans released from glycoproteins were analyzed on an Axima MALDI Resonance mass spectrometer with a QIT-TOF configuration (Shimadzu). A nitrogen laser was used to irradiate samples at 337 nm, with an average of 200 shots accumulated. Permethylated glycan standards and *N*-glycans from glycoproteins dissolved in methanol were applied to a μfocus MALDI plate target (900 μm, 384 circles, HST). A matrix solution (0.5 μL) of 2,5-dihydroxybenzoic acid (20 mg/mL) in a mixture of methanol/water (1:1) containing 0.1 % trifluoroacetic acid and 1 mM NaCl was added to the plate and mixed with samples. The mixture was air dried at room temperature before analysis.

Among the four different resolution settings of the instrument FWHM 70, 250, 500 and 1000 for precursor isolation, the window at FWHM 500 with a width of 3–5 mass units was considered appropriate (Supplementary Results) and used for the present study. This conclusion was based on the initial testing using two permethylated standard *N*-glycans Hybrid-Octa (MNa^+ *m/z* 1824) and FA2 (MNa^+ *m/z* 1835) with a mass difference of 11 (Supplementary Results).

Product-ion spectra acquired at each stage were introduced into the GIPS-II program as a mzXML file (Shimadzu) for peak identification, using a signal-to-noise ratio 3:1 as the filtering parameter. Candidate *N*-glycans were extracted from CarbBank and stored in the GIPS-II system. The probability and DP values of the peaks were manually fed back to the mass spectrometer data system, and to instruct further product-ion scanning using the fragment ion with the highest DP to generate MS^n spectrum. If the probability of one candidate exceeded the pre-defined threshold 0.70, the process stopped and reported the branching patterns. If the fragments of one candidate were almost covered by those of other candidates, the DP_{biased} of all the remained peaks was calculated by the biased-DP-model. The peak with the highest DP_{biased} was selected and extracted as precursor to generate the next stage MS spectrum until unique peaks of one candidate glycan were observed.

For MS^n using the MIPS approach (Ashline et al., 2007), the most intense peak was selected as the precursor for the next-stage MS scanning whereas for MS^n using the GIPS-II approach, the peak with the highest DP or DP_{biased} was selected as the precursor as directed by the computer program.

2.5. Capillary electrophoresis (CE)

Capillary electrophoresis was performed using Beckman PA 800 plus CE system equipped with a laser-induced fluorescence detector. A fused-silica capillary with a total length of 50.2 cm was used as the separation channel, and the inner diameter was 50 μm. The new capillary was sequentially conditioned with 1.0 M NaOH, 0.1 M NaOH, water and *N*-glycan separation buffer (background buffer BGE, Beckmann) for 10 min respectively. The standard and released *N*-glycans derivatized with 8-aminopyrene-1,3,6-trisulfonic acid (APTS, Sigma-Aldrich) were injected into the capillary for 10 s at 3.45 kPa (0.5 psi), and the end of the capillary was washed with the separation buffer to remove attached samples. A negative voltage of 25 kV was applied, and the system was performed in reversed polarity.

3. Results and discussion

For the present work of *N*-glycan branching pattern analysis, only *N*-glycans listed in the CarbBank database are selected to build the library. All the *N*-glycans with identical molecular mass as the sample's (± 0.05 Da), identified from the primary mass spectrum, are extracted from the library and considered as candidate glycans. Glycan identification is then achieved by assigning the sample glycan from the listed candidate glycans based on comparison of the acquired sample spectrum and the candidates' theoretical spectrum at each stage of MS^n. For construction of theoretical spectrum, three glycosidic cleavages were considered which are necessary to give sufficient fragmentation while maintaining specificity.

3.1. N-Glycan analysis using GIPS

High mannose *N*-glycans Man-7D1 and Man-7D3 (Table 1) are identical in composition but different in branching patterns with one of the non-reducing terminal mannose residues at a different branch (D1 and D3, respectively). Here, the two isomeric standard *N*-glycans were used as representatives to demonstrate the identification process using DP-model of the GIPS approach (Sun et al., 2018).

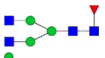
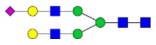
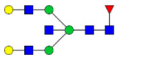
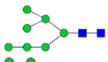
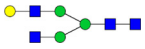
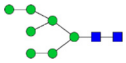
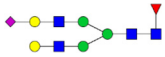
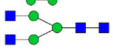
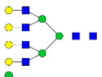
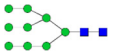
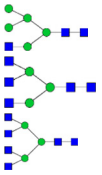
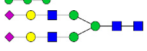
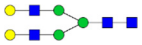
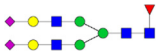
Based on the primary mass spectrum with MNa^+ at *m/z* 1988.0, a total of 8 candidates were extracted (G_1 to G_8, Fig. 1, bottom), and each of the candidates was assigned with an equal probability 0.125 ($1/8 \approx 0.13$) in both cases of Man-7D1 and Man-7D3.

The MS^2 spectrum of Man-7D1 obtained from MNa^+ led to recalculated candidate probabilities: those of G_1, G_3, G_5 and G_8 increased to 0.23 while others decreased to either 0.08 or 0 (Fig. 1a). Among the fragment ions in the spectrum of MS^2, *m/z* 1084.5 was identified to have the highest DP (1.03) for use as the precursor to produce a product-ion spectrum which potentially has the most distinctive feature. When the MS^3 spectrum was obtained, further calculated probabilities indicated a significantly high probability for G_1 (0.94), and therefore Man-7D1 (G_1) was correctly reported as the actual glycan.

The identity of the Man-7D3 was similarly assigned. Although Man-7D3 gave a different MS^2 spectrum from that of Man-7D1, making it readily differentiated from Man-7D1 (Fig. 1b), more rounds of scanning were required to differentiate it from the other 6 candidates (G_3–G_6). GIPS selected *m/z* 1492.7 (DP 1.36) in the spectrum of MS^2 to carry out MS^3, and *m/z* 871.4 (DP 1.03) in the spectrum of MS^2 to carry out another MS^3. Thus after 3 and 4 rounds of scanning the identities of Man-7D1 and Man-7D3 were successfully determined (Fig. 1).

Using the similar process, *N*-glycans A2G2S2, Hybrid-Octa, A3, A4, Man-6, A2, A2G1, FA2G2, Man-9, A2G2S1, FA2BG2, FA2G2S1, A4G4 and FA2G2S2 (Table 1) were successfully identified based on DP for selection of the precursor ions (Figures S1-S14). The peak selection pathway and corresponding probability for each *N*-glycan were shown in Table 2, and detailed identification results were described in Supplementary Results.

Table 1
Standard *N*-Glycan branching patterns identified by the GIPS-II approach.

Glycan	Mol Mass (*Dalton*)	Branching pattern	Glycan	Mol Mass (*Dalton*)	Branching pattern
FA2	1462.5		A2G2S1	1931.7	
Man-6	1396.5		FA2BG2	1989.7	
Man-7D1	1558.5		A2G1	1478.5	
Man-7D3	1558.5		FA2G2S1	2077.7	
A2	1316.5		A4G4	2370.9	
Man-9	1882.6		Hybrid-Octa	1437.5	
A2G2S2	2222.8		A3	1519.6	
A2G2	1640.6		A4	1722.6	
FA2G2	1786.7		FA2G2S2	2368.8	
A3G3S3	2879.0		–	–	–

Glycan nomenclature: Green circle, Mannose; Yellow circle, Galactose; Red triangle, Fucose; Blue square, *N*-acetyl glucosamine; Yellow square, GalNAc; Green square, ManNAc; Purple diamond, *N*-acetylneuraminic acid; White diamond, *N*-glycolylneuraminic acid. Trivial names are used for the *N*-glycans, and Hybrid-Octa is used for the hybrid *N*-glycan octasaccharide. The linkage and the 3- and 6-branch of the N-glycan mannose core were not specified.

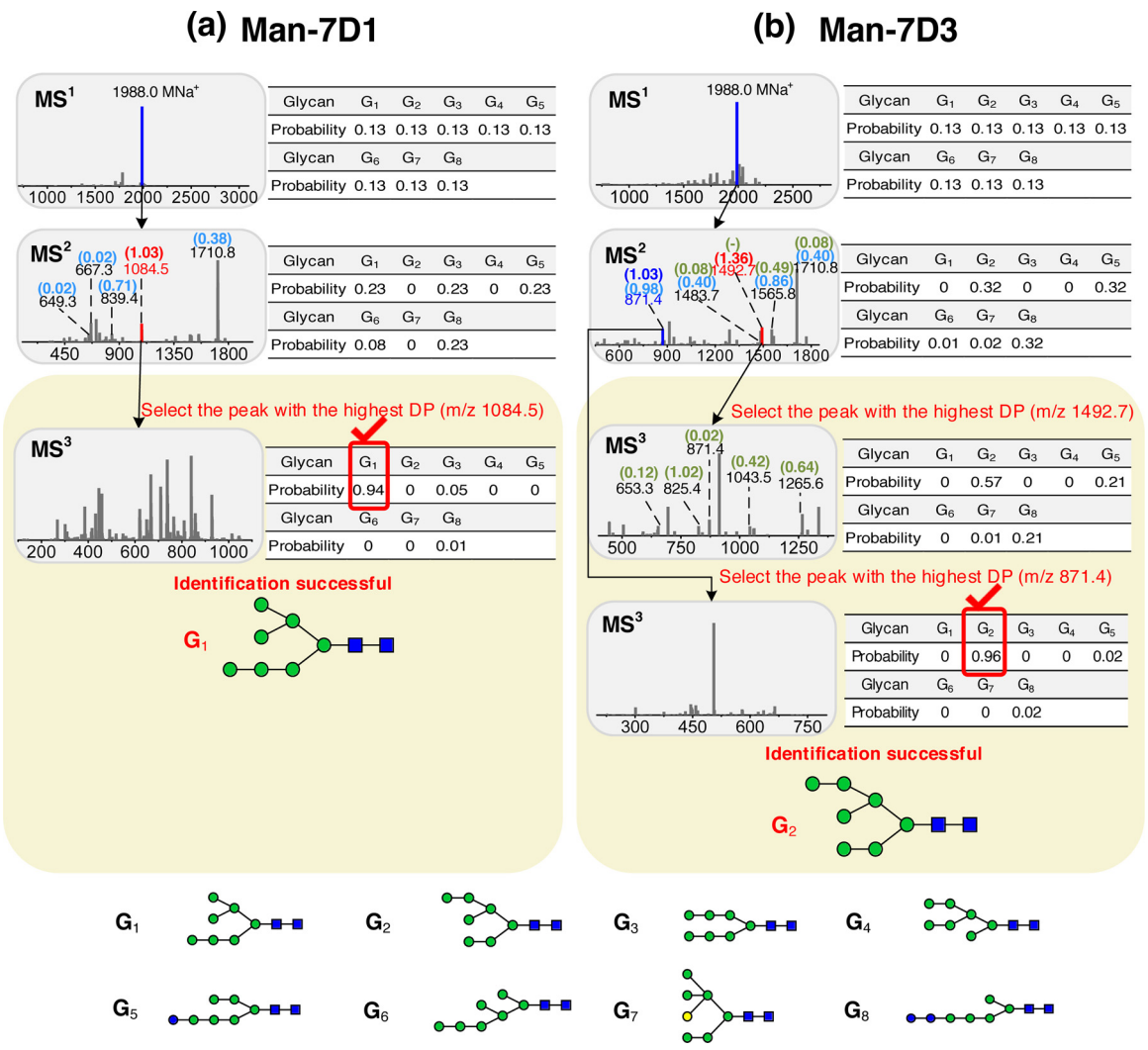


Fig. 1. Identification process of the isomeric pair Man-7D1 and Man-7D3 using GIPS.
MS^3 was required for the identification of Man-7D1 (a), while two MS^3 were required for Man-7D3 (b) from the 8 candidate glycans G_1 to G_8. Mass spectrometry process is shown in the left column and the calculated probabilities are in the right column. The fragment ions in red were selected as the precursors for the next round of product-ion scanning. The calculated DPs are in brackets above the *m/z* values in the spectrum.

3.2. The challenge for GIPS

For the *N*-glycan A2G2 (Table 1 and Fig. 2), among the 7 candidates G_1 – G_7, the probabilities of G_1 and G_2 are the highest and identical, but they remained at 0.50 or below even after five consecutive scanning, including two MS^3 (MS^3-a and MS^3-b) using two different fragments (*m/z* 1329.6 and *m/z* 1143.6) in the MS^2 spectrum as the precursors and two MS^4 (MS^4-a and MS^4-b) using two different fragments (*m/z* 939.5 and *m/z* 866.4) in the MS^3-b spectrum as the precursors, while the probabilities of G_3 – G_7 decreased to 0. These values indicated that the sample *N*-glycan is unlikely to be G_3 – G_7. However, although it is likely to be either G_1 or G_2 there were no discriminative fragments in the five product-ion spectra acquired including one MS^2, two MS^3 and two MS^4 (Fig. 2). Based on the DP using GIPS, no further information can be obtained for distinguishing G_1 and G_2.

It was noticed that the calculated theoretical distinctive fragments of G_1 are lacking or much less than those of G_2, i.e. almost all the theoretical fragments of G_1 are identical to those of G_2. If the sample glycan is G_2 it can be readily identified using the DP-model of GIPS. However, in the case that the sample is G_1, selection of any further fragment ions with high DP value cannot produce any characteristic product-ion spectrum for distinguishing G_1 from G_2 (see Supplementary Results).

A new methodology is required to solve the problem in the latter case, and we designed two additional approaches, based on the statistical perspective, to add into the GIPS and these formed the improved strategy, GIPS-II.

3.3. The concept of GIPS-II

The main purpose for precursor ion selection is to produce a product-ion spectrum which contains distinctive fragment peaks and can be used to assign the sample glycan. In the GIPS approach, a fragment ion with the highest DP is selected as the precursor for the next round of scanning as it is most likely to produce a product-ion spectrum containing distinctive fragmentation for use to assign the sample glycan from the list of candidate structures. However, as shown above, GIPS may fail the identification in the following circumstances. First, the sample *N*-glycan does not have any distinctive fragments compared to other candidates. Second, the sample glycan may have a few characteristic fragment peaks but the number of such peaks is less than those of other candidate glycans, such as in the case of *N*-glycan NA2 discussed above, in which the sample glycan is G_1 but all or most of its fragment ions are shared with candidate G_2, i.e. there is no or fewer specific fragments to be used for differentiation of G_1 from G_2. In this case, the calculated probability value of each candidate cannot be

Table 2
Peak selection pathway of standard *N*-glycans and calculated results.

Glycans	MNa^+ (m/z)	Number of Candidates	GIPS	GIPS-II	
			Peak selection pathway and Probability using DP-model	p_{BF}	Selected peaks for testing biased-DP-model
A2G2S2	2792.4	1	MS2: 2792.4 (1.00)		
Hybrid-Octa	1824.9	3	MS2: 1824.9 (0.72)		
A3	1907.0	4	MS2: 1907.0 (0.41)		
			MS3: 1907.0→1111.5 (0.83)		
A4	2152.1	5	MS2: 2152.1 (0.23)		
			MS3: 2152.1→1615.8 (0.81)		
Man-6	1783.9	5	MS2: 1783.9 (0.33)		
			MS3: 1783.9→1084.5 (0.66)		
			MS4: 1783.9→1084.5→667.3 (0.70)		
Man-7D3	1988.0	8	MS2: 1988.0 (0.32)		
			MS3: 1988.0→1492.7 (0.57)		
			MS3: 1988.0→871.4 (0.96)		
Man-7D1	1988.0	8	MS2: 1988.0 (0.23)		
			MS3: 1988.0→1084.5 (0.94)		
A2	1661.8	4	MS2: 1661.8 (0.69)		
			MS3: 1661.8→662.3 (0.87)		
FA2	1835.9	5	MS2: 1835.9 (0.48)	699857.8	1835.9→1317.6
			MS3: 1835.9→1576.8 (0.49)		
			MS4: 1835.9→1576.8→662.3 (0.50)		
			MS3: 1835.9→662.3 (0.50)		
			MS4: 1835.9→662.3→458.2 (0.50)		
A2G1	1865.9	5	MS2: 1865.9 (0.69)		
			MS3: 1865.9→1402.7 (0.73)		
A2G2	2070.0	7	MS2: 2070.0 (0.43)	22.2	2070.0→1606.8
			MS3: 2070.0→1329.6 (0.49)		
			MS3: 2070.0→1143.6 (0.50)		
			MS4: 2070.0→1143.6→939.5 (0.50)		
			MS4: 2070.0→1143.6→866.4 (0.50)		
FA2G2	2244.1	9	MS2: 2244.1 (0.28)		
			MS3: 2244.1→1329.6 (0.31)		
			MS3: 2244.1→1780.9 (0.93)		
Man-9	2396.2	12	MS2: 2396.2 (0.32)		
			MS3: 2396.2→1696.8 (0.77)		
A2G2S1	2431.2	2	MS2: 2431.2 (0.95)		
FA2BG2	2489.3	4	MS2: 2489.3 (0.37)		
			MS3: 2489.3→1766.9 (0.61)		
			MS3: 2489.3→1315.6 (0.76)		
FA2G2S1	2605.3	2	MS2: 2605.3 (0.95)		
FA2G2S2	2966.5	2	MS2: 2966.5 (0.73)		
A4G4	2968.5	4	MS2: 2968.5 (0.25)		
			MS3: 2968.5→1578.8 (0.76)		
A3G3S3	3602.8	5	MS2: 3602.8 (0.31)	2272.2	3602.8→3227.6
			MS3: 3602.8→1129.5 (0.32)		
			MS3: 3602.8→2403.2 (0.41)		
			MS4: 3602.8→2403.2→852.4 (0.41)		
			MS4: 3602.8→2403.2→662.3 (0.41)		

beyond 0.50 with the DP-model used in GIPS. This difficult situation is considered here as "coverage curse" for G_1.

The improved GIPS-II was designed to break the "coverage curse" for G_1, and assign the branching pattern of the sample *N*-glycan when the probabilities of the two most likely candidates remain equal (for example at 0.50) after repeated product-ion scanning.

Intuitively, one mass spectrum can be explained by either candidate G_1 and G_2 with the same calculated probability, and the theoretical fragments of G_1 are covered or nearly covered by those of candidate G_2. G_2 is unlikely to be the sample *N*-glycan if distinctive peaks of G_2 were not observed in the acquired mass spectrum after enough rounds of scanning (in the present study, we restrict five rounds of scanning after MS^1). Based on this principle, G_2 can be excluded and the sample *N*-glycan can be tentatively assumed as G_1 before the hypothesis is theoretically and experimentally confirmed by GIPS-II.

3.4. Implementation of GIPS-II

After a few rounds of scanning, several candidates carve up all the probability, and fragments of one candidate (without loss of generality, we call it G_1) are covered by others, and the probability of each candidate remains at or below 0.50. In this case, GIPS-II makes a *hypothesis* that the sample *N*-glycan is G_1. This is followed by a *significance test* using the BayesFactor-model (BF-model) to statistically *confirm* the hypothesis. A new peak selection strategy, the *biased-DP*-model, is designed to experimentally verify the hypothesis. The biased-DP-model selects fragment ions as precursors which have the potential to produce most distinctive peaks of the sample *N*-glycan for experimental confirmation in the case of "coverage curse" (Fig. 3).

3.4.1. Significance test to judge whether the G_1 is the N-glycan sample

The theoretical spectrum of every candidate is closely associated with the experimental spectrum at each stage of MS^n, based on which the theoretical spectrum was calculated. For the purpose of making explanations simpler, only two candidates, G_1 and G_2, are considered here. Theoretical spectrum of G_1 is viewed as a set of fragment masses $S_1 = \{m_1^1, m_2^1, \cdots, m_p^1\}$, and the theoretical spectrum of G_2 as $S_2 = \{m_1^2, m_2^2, \cdots, m_q^2\}$. Furthermore, the intersection of these two sets is $S_1 \cap S_2 = \{m_1^{1,2}, m_2^{1,2}, \cdots, m_k^{1,2}\}$. Experimental spectrum can be regarded to be a sampling process, and each peak in the spectrum is considered to

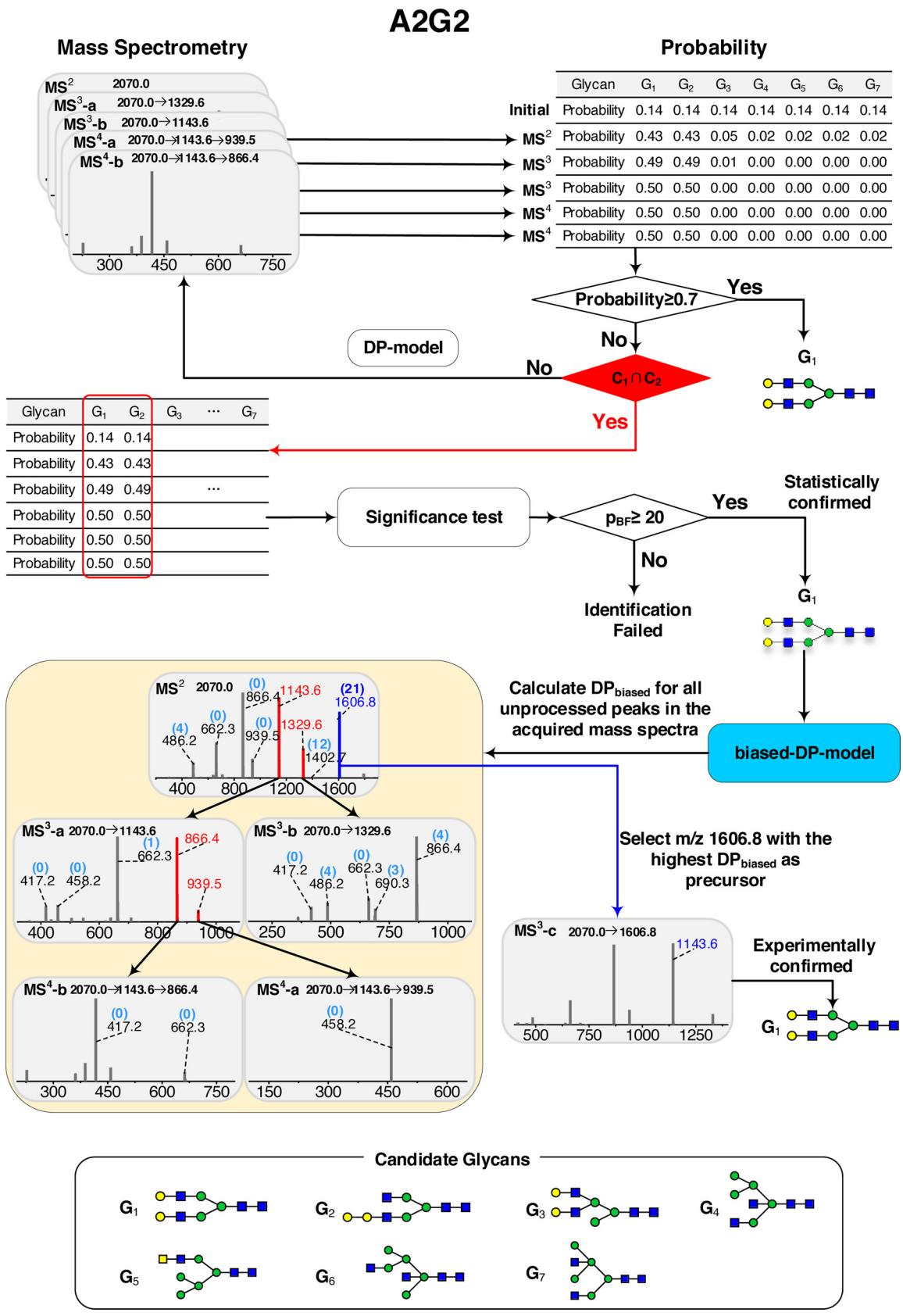


Fig. 2. Identification process of A2G2 using GIPS-II.
Seven candidate glycans G_1 to G_7 were extracted from the database, and candidate glycans G_1 and G_2 were calculated with the identical highest probability using DP-model. MS^3 was required using GIPS-II with the significant test and biased-DP-model to identify A2G2. The fragment ions in red were selected as precursors using DP-model, and fragment ions in blue were selected using biased-DP-model as the precursors for the next round of product-ion scanning. The calculated DP_{biased} values are in brackets above the *m/z* values in the spectrum. C_1 and C_2 represent two different conditions; C_1: the number of used spectra ≥ 5; C_2: the used spectra have no more information to distinguish between G_1 and G_2.

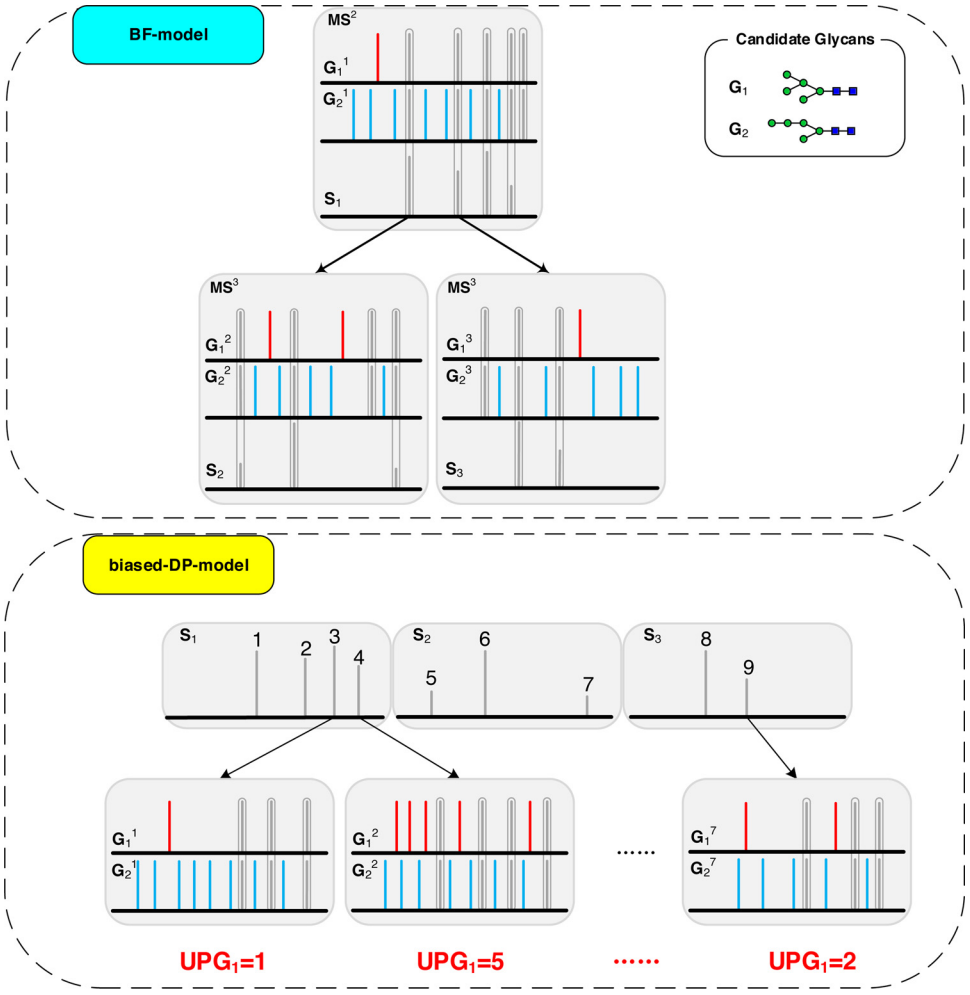


Fig. 3. The insights of BF-model and biased-DP-model.
In BF-model, S_i is the i-th experimental spectrum. G_1^i and G_2^i represent the i-th theoretical spectra of the two respective candidates. In biased-DP-model, S_i is the i-th experimental spectrum. G_1^i and G_2^i represent the i-th theoretical spectra of the two respective candidates. In each spectrum, the red line denotes the unique peak of G_1 and the blue line indicates the unique peak of G_2, and the grey line represents the common peak of G_1 and G_2. UPG_1 denotes the number of unique peaks of G_1. The biased-DP-model selects the peak in S_i whose theoretical product-ion spectrum provides the most unique peaks of G_1.

be a random sampling from one of the two sampling space, S_1 or S_2. When G_1 is the sample *N*-glycan, the sampling space is S_1, otherwise is S_2.

If distinctive peaks for G_1 and G_2 are not observed even after acquisition of several successive experimental mass spectra, the candidate (G_2) with many distinctive peaks is unlikely to be the actual glycan, and therefore the sample *N*-glycan is more likely to be G_1. In the present work, we turned to Bayes factor for significance test to determine if G_1 has enough confidence to be the sample *N*-glycan.

Before the significance test, there are already several acquired mass spectra. Each acquired mass spectrum is coming with two sampling space: S_1 from the theoretical spectrum of G_1 and S_2 from the theoretical spectrum of G_2. One peak in the spectrum can be regarded as one time sampling from one of the two sampling spaces. The intersection of these two sets is denoted as $S_1 \cap S_2$. Each experimental mass spectrum E^i can be denoted as a quadruple, and all experimental spectra can be denoted as

$$\{E^1, S_1^1, S_2^1, S_1^1 \cap S_2^1\}, \{E^2, S_1^2, S_2^2, S_1^2 \cap S_2^2\}, \cdots, \{E^n, S_1^n, S_2^n, S_1^n \cap S_2^n\}$$

$$p(spec|C_1) = \prod_{i=1}^{n} \frac{C_{|S_1^i \cap S_2^i|}^{|E^i|}}{C_{|S_1^i|}^{|E^i|}} \quad (1)$$

$$p_{BF} = \frac{p(C_1|spec)}{p(C_2|spec)} = \frac{p(spec|C_1) \times p(C_1)}{p(spec|C_2) \times p(C_2)} = \prod_{i=1}^{n} \frac{C_{|S_2^i|}^{|E^i|}}{C_{|S_1^i|}^{|E^i|}} \quad (2)$$

Two prior odds are deemed equally before the data are collected, and the posterior odds are equal to the value of Bayes factor, noting as $p_{BF} = \frac{p(spec \mid C_1)}{p(spec \mid C_2)}$ (Eqs. (1) and (2)). It is clear that the p_{BF} plays a crucial role in establishing the relative evidential support for the null and alternative hypotheses.

A peak in the experimental spectra can be regarded as a sampling from the sampling space, and all peaks mean many times of random sampling and if the result satisfies: i) At least 5 consecutive mass spectra were used for the calculation of probability of candidate G_1 and G_2; ii) The probability values of the top 2 candidates remained the same. Entropy of the probability system also remained the same after a series of spectra were produced, and the following produced spectra has little potential to pick the sample *N*-glycan out.

3.4.2. Further experiment to confirm the sample N-glycan in this particular situation

In order to confirm if the candidate G_1 is the sample *N*-glycan, biased-DP-model selects the peak whose product-ion spectrum is promising to produce more unique peaks of G_1. As the BF-model for significance test has confirmed that the candidate G_1 is the sample *N*-glycan, the remaining step is validation using the new peak selection strategy based on biased-DP-model. The peak selection strategy based on the DP_{biased} value is a little tricky. Unlike the DP-model giving different weights to distinctive peaks of different candidates according to the latest probability of every candidate, we give the whole weight to distinctive peaks of G_1 and 0 to distinctive peaks of other candidates. The value of DP_{biased} was calculated using Eq. (3), where n is the number of candidates, and w_i is the weight of the i-th candidate for the calculation of DP_{biased}, and the sum of all w_i is equal to 1, and the d_i represents the count of distinctive peaks for the i-th candidate. As

mentioned above, when calculating DP_{biased}, the weight of G_1 (w_1) is set as 1 and weights of all other candidates ($w_2 to w_n$) are set as 0. Therefore, the value of DP_{biased} is equal to the count of distinctive peaks for G_1. The peak whose theoretical product-ion spectrum provides more distinctive peaks for G_1 will be selected for next stage MS scanning. This strategy has the capability to validate the sample glycan rapidly.

$$DP_{biased} = \sum_{i=1}^{n} (w_i \times d_i) \quad (3)$$

In the rare and very special cases, the fragments of glycan G_1 are totally covered by those of G_2, which means G_1 has no distinctive peaks, the experimental confirmation cannot be achieved. In this situation, our experiment will stop after significance test.

3.5. Analysis of individual N-glycans using GIPS-II

As discussed above, for analysis of the *N*-glycan A2G2, even after five consecutive scanning including two MS^3 (MS^3-a and MS^3-b) and two MS^4 (MS^4-a and MS^4-b), it was not possible to distinguish between G_1 and G_2. We now illustrate how this problem can be solved using GIPS-II.

As the calculated theoretical fragments of G_1 were covered or nearly covered by those of candidate G_2, hypothesis was made to assume G_1 was the sample *N*-glycan based on the accounted fragments coverage. Significance test was then applied, and the value of p_{BF} was calculated to be 22.2, greater than the predefined threshold 20. Therefore, the candidate G_1 with less distinctive peaks was statistically confirmed to be the sample *N*-glycan. DP_{biased} values of all the remaining peaks in the acquired five experimental spectra were then calculated in order to select the best precursor to carry out the final product-ion scanning for experimental confirmation. As shown in Fig. 2, the peak at *m/z* 1606.8 was calculated to have the highest DP_{biased} value, indicating that most distinctive peaks of G_1 will be produced in the product-ion spectrum of *m/z* 1606.8. Therefore, *m/z* 1606.8 was selected as the precursor to generate a new MS^3 spectrum. As shown in Fig. 2, a new distinctive ion *m/z* 1143.6 for G_1 was observed, and the sample *N*-glycan was thus identified as G_1. The glycans FA2 and A3G3S3 (Table 1) were similarly determined using the significance test and biased-DP-model (Figures S15 and S16, Table 2).

We compared GIPS-II with the widely-used MIPS strategy for the identification of 13 representative *N*-glycans including complex type, high-mannose type and Hybrid type. All *N*-glycans were successfully identified by GIPS-II before MS^4 scanning, whereas more than half of these *N*-glycans could not identified using MIPS even after MS^5 scanning. The results were described in Supplementary Results and Table S1.

3.6. Application to N-glycan profiling of Glycoproteins

Having demonstrated the identification of individual *N*-glycans, we further applied GIPS-II to *N*-glycan profiling and subsequent detailed identification of individual components of glycoproteins RNase B (Fig. 4a) and the therapeutic mAb adalimumab (Fig. 5a), used initially as samples containing only relatively simple *N*-glycans. The MS^1 spectrum of each glycoprotein exhibited multiple components. Unlike the examples presented above, all the ions present in the primary spectrum were required for product-ion scanning to determine their identities. The results obtained from GIPS-II at this development stage were checked manually, corroborated by other method, or compared with literature data.

As shown in Table 3, for the sample isolated from RNase B, product-ion scanning for each of the five MNa^+, *m/z* 1579.8, 1783.9, 1988.0, 2192.1 and 2396.2, was performed. Among these, *m/z* 1579.8, 1783.9 and 2396.2 were identified as Man-5D1, Man-6 and Man-9 (Table 3), with the probabilities 0.88, 0.94, and 0.94, respectively (Table S2),

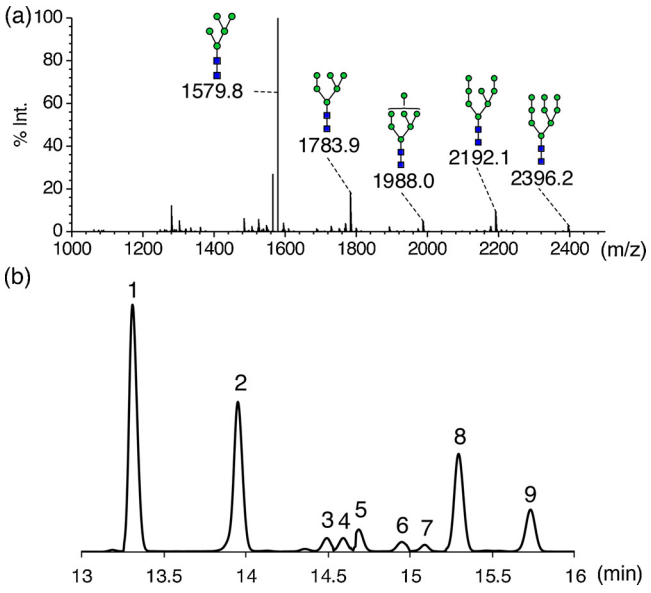


Fig. 4. Profiling and subsequent assignment of *N*-glycans released from RNase B.
(a) MALDI mass spectrum of the released *N*-glycans as permethylated derivatives contained five main components with MNa^+ at *m/z* 1579.8, 1783.9, 1988.0, 2192.1 and 2396.2, respectively. MS^n with GIPS-II identified these as high-mannose type *N*-glycans containing five, six, seven, eight and nine mannose residues, respectively. The matched structures in the database were shown, but the linkage and the 3- and 6-branch of the N-glycan mannose core were not specified. **(b)** Electropherogram of the released *N*-glycans from RNase B.

within three rounds of scanning. The assignment was confirmed by CE analysis of *N*-glycan standards (Fig. 4b). Also revealed by CE, Man-5, Man-6 and Man-9 each contained a single component, and therefore could be readily assigned by DP-model. The results are in agreement with the previous studies (Morris, Peterson, & Tarlov, 2009; Valk-Weeber, Dijkhuizen, & van Leeuwen, 2019).

However, multiple isomeric components were present in Man-7 (Table 1, *m/z* 1988.0, and peaks 3–5 in the electropherogram, Fig. 4b) and Man-8 (Table 1, *m/z* 2192.1, and peaks 6–8, Fig. 4b). It is interesting that a successful result was obtained for Man-8 to be Man-8D1D3 (Table 3) as it has a ~85 % purity (peak 8, Fig. 4b). The identified Man-8D1D3 has been previously reported to be the main isomeric structure of Man-8 in RNase B (Morris et al., 2009; Valk-Weeber et al., 2019). Man-7 contained three isomers, Man-7D1, Man-7D2, and Man-7D3 with approximate concentrations of 26 %, 26 %, and 48 %, respectively (peaks 3–5, Fig. 4b). Using the present approach, Man-7D2 and/or Man-7D3 can be identified as the main Man-7 component (with a probability 0.77). The combined concentration of these two, based on electrophoresis (Fig. 4b), is 74 % (26 % + 48 %). As the fragment ions of these two are identical the two isomers are indistinguishable.

For the sample isolated from adalimumab, we identified 12 glycans including an isomeric pair (Table 3), among which 11 *N*-glycans were successfully identified using the DP-model. However, the assignment of *N*-glycan at *m/z* 2401.2, required the significance test and DP_{biased}. The detailed identification process for RNase B and adalimumab was shown in Table S2. The identified structures were essentially in agreement with CE analysis (Fig. 5b) and previous report on the Fc-glycosylation profile (Reusch et al., 2015).

3.7. Application to N-glycan profiling of a human plasma glycoprotein and human serum

Altered glycosylation patterns of plasma proteins are associated with autoimmune disorders and pathogenesis of various cancers (Ren

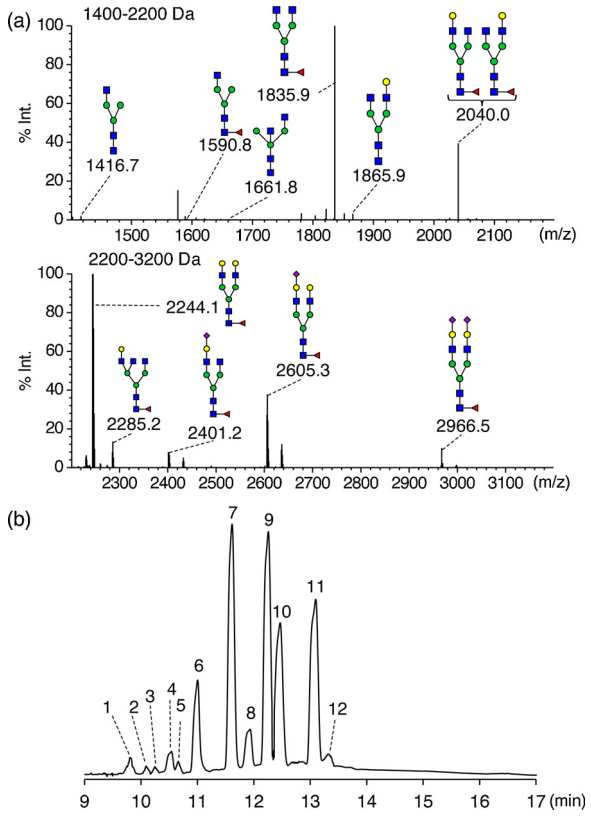


Fig. 5. Profiling and subsequent assignment of *N*-glycans released from mAb adalimumab.

(a) MALDI mass spectrum of the released *N*-glycans as permethylated derivatives showed twelve *N*-glycans with MNa^+ at *m/z* 1416.7, 1590.8, 1661.8, 1835.9, 1865.9, 2040.0, 2244.1, 2285.2, 2401.2, 2605.3, and 2966.5, respectively. The structures identified by GIPS-II are shown next to the ion peaks. The peak at *m/z* 2040.0 is a mixture of two isomers (peaks 9 and 10, in panel b). The matched structures in the database were shown, but the linkage and the 3- and 6-branch of the N-glycan mannose core were not specified.

(b) Electropherogram of the released N-glycans indicated a total of twelve components, among these the identities of six can be corroborated by available N-glycan standards (peak 1, FA2G2S2; peak 6, FA2G2S1; peak 7, FA2; peak 9, FA2G1-a; peak 10, FA2G1-b; peak 11, FA2G2) while others by literature data (Reusch et al., 2015).

et al., 2016). AGP is the high-abundance acute-phase plasma protein, and is proposed to deliver bound drugs to cells by conformational changes induced by the cell membrane (Kishino et al., 2002). As a result, alternations of glycan structures, especially detailed branching patterns, should have significant influence on the drug binding capacity.

One of the important aspects of *N*-glycan identification is the location of the fucose residue. It can form core-fucosylated structures or Lewis x epitopes, which are two important functional structures for glycoproteins. As shown in Fig. 6, ten *N*-glycans were observed at *m/z* 2070.0, 2244.1, 2519.3, 2693.4, 2968.5, 3142.5, 3316.6, 3417.7, 3490.7, and 3866.9 in the primary mass spectrum, and all were successfully identified (Tables S2 and S3). GIPS-II, based on the significance test and biased-DP-model, was required for the assignment of *m/z* 2070.0 but the rest only needed the DP-model. Among these *N*-glycans, five were fucosylated and the fucose residues were identified unambiguously to be attached to the nonreducing terminal GlcNAc forming Lewis x epitopes rather than the reducing core GlcNAc. The results are in agreement with the report that AGP has five glycosites occupied by complex type *N*-glycans with Lewis x epitopes (Fournier, Medioubi-N, & Porquet, 2000; Yazawa et al., 2016; Wu, Struwe, Harvey, Ferguson, & Robinson, 2018).

Table 3
Application of GIPS-II to N-Glycan Profiling and Subsequent Assignment.

Glycoprotein	MNa^+ (*m/z*)	Number of Candidates	Assignment	CE Peak No[a]
RNase B	1579.8	3	Man-5D1	1
	1783.9	5	Man-6	2
	1988.0	8	Man-7D2/7D3	3,5
	2192.1	9	Man-8D1D3	8
	2396.2	7	Man-9	9
mAb[b]	1416.7	2	A1	2
adalimumab	1590.8	2	FA1	4
	1661.8	4	A1B	5
	1835.9	5	FA2	7
	1865.9	5	A2G1	8
	2040.0	5	FA2G1-a	9
	2040.0	5	FA2G1-b	10
	2244.1	9	FA2G2	11
	2285.2	7	FA3G1	12
	2401.2	2	FA2G1S1	3
	2605.3	2	FA2G2S1	6
	2966.5	2	FA2G2S2	1

[a] Corroborated by CE analysis (Figs. 4b and 5 b) with peak numbers shown in the Table.

[b] Predicted. For the *N*-glycans isolated from mAb adalimumab, 12 were assigned, and 6 of which were corroborated by CE analysis (Fig. 5b) with available glycan standards (1: FA2G2S2, 6: FA2G2S1, 7: FA2, 9: FA2G1-a, 10: FA2G1-b, 11: FA2G2).

We also applied the strategy to analysis of human serum *N*-glycans, including the initial profiling and the subsequent branching pattern assignment of individual *N*-glycan components (Fig. 7). For this complex sample, we identified 45 *N*-glycans of the high-mannose-, complex-, and hybrid-types, among which 42 have been variously reported (Fig. 7, Table S4 and references therein). In these *N*-glycans, 43 *N*-glycans were successfully identified with specific branching patterns using the DP-model, but two *N*-glycans (*m/z* 2315.2 and 3602.8) could not be determined even after six rounds of MS scanning (Table S2). Hypothesis was then made to assume one of the remaining two candidates was the sample glycan before significance test and biased-DP-model were utilized to statistically and experimentally confirm the structure. The branching patterns of these two *N*-glycans were successfully identified as asialo-digalacto-triantennary (*m/z* 2315.2) and fully sialylated tri-antennary (*m/z* 3602.8) structures, respectively.

4. Conclusions

MS^n has been important for sequence determination of biomolecules. For sequencing of peptide, precursor ion selection is straightforward and the most intense ions are suitable to generate useful product-ion spectra as protein/peptide sequence is linear. Multiple branching is an important feature of glycans. The high-mannose, complex and hybrid types of *N*-glycans are all of different branching patterns, and within each type there are also isomeric sequences of high similarity. For the multiply branched glycans, it is not always easy to select the best precursor ion for the next round scanning. Selecting the most intense ion as the precursor may not produce structurally informative spectrum.

In our previous work, GIPS was developed for the first time for intelligent selection of precursors to produce a product-ion spectrum which contains the most distinctive fragment ions for identification of the branching patterns. However, distinctive fragments of some glycans, particularly *N*-glycans, may be much fewer or do not occur at all, and this may lead to failure in glycan identification by GIPS. We have now introduced a hypothesis and a significant test procedure based on the BF-model, and a new peak selection method using biased-DP-model. The improved strategy has been successfully used in identification of 17 individual *N*-glycans. The approach was also used to determine *N*-

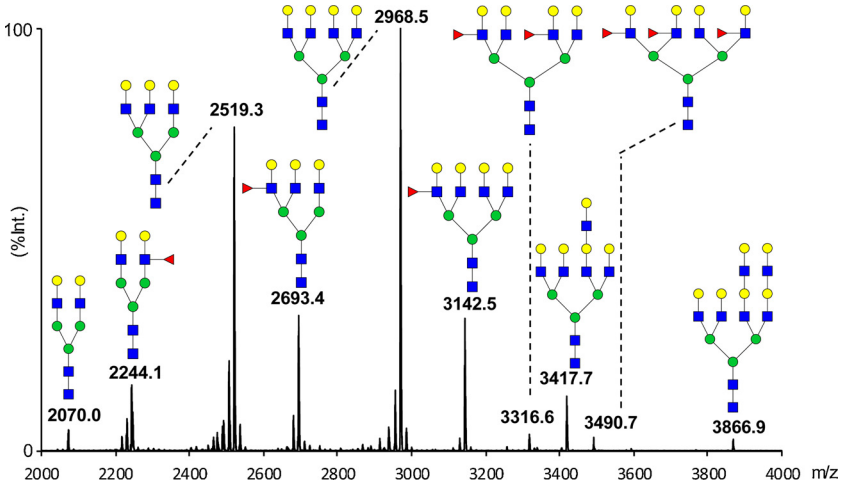


Fig. 6. Profiling and the subsequent assignment of individual *N*-glycans released from human asialo-AGP using GIPS-II.
Positive-ion MALDI spectrum of *N*-glycans from asialo-AGP after permethylation. The annotation of the ions was made based on GIPS-II using either the DP-model or the biased-DP-model. The detailed results of the assignment are shown in Table S2. The matched structures in the database were shown, but the linkage and the 3- and 6-branch of the N-glycan mannose core were not specified.

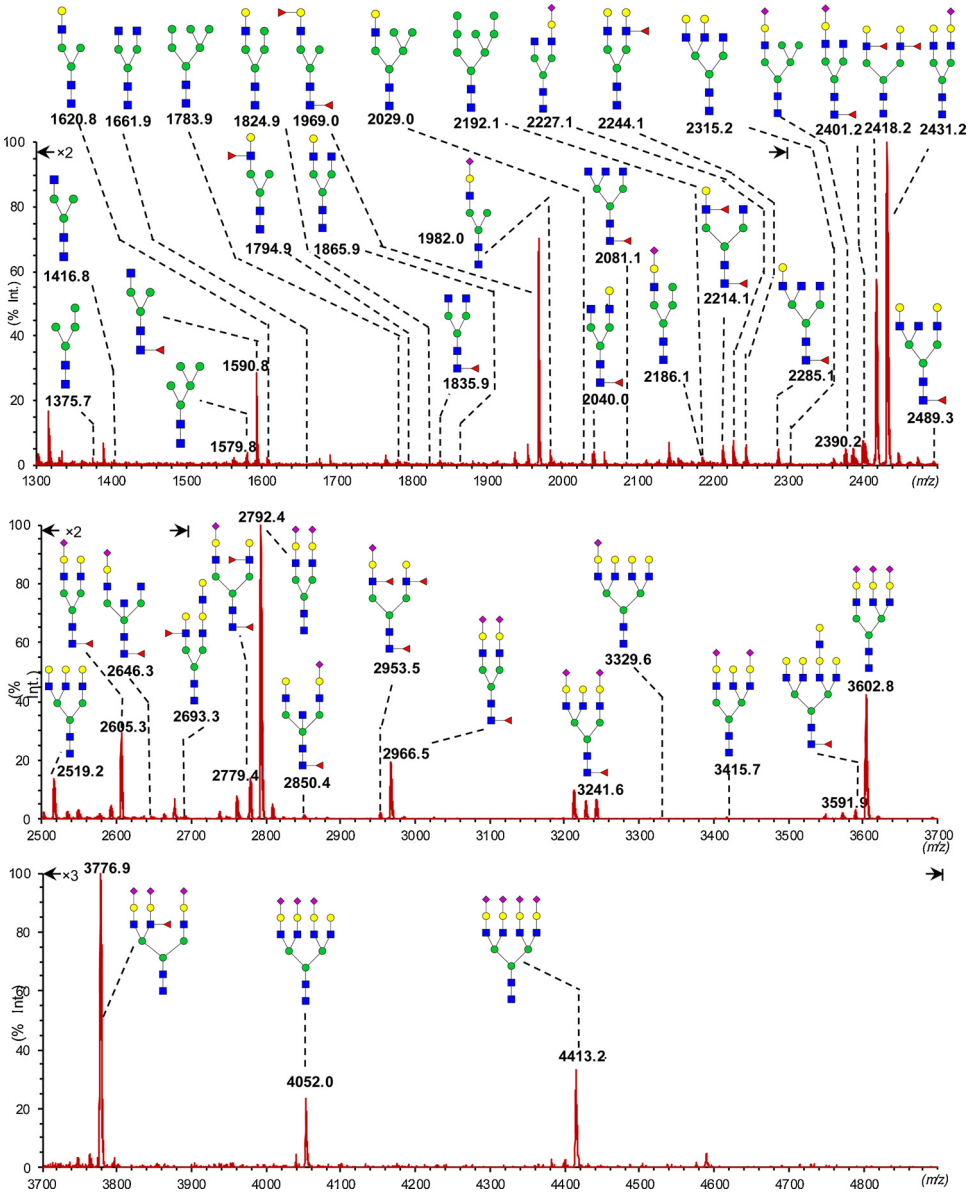


Fig. 7. *N*-glycan profiling and branching pattern assignment of a human serum sample by GIPS-II.
Positive-ion MALDI spectrum of *N*-glycans from human serum after permethylation. The annotation of the ions was based on GIPS-II using either the DP-model or the biased-DP-model. The detailed results of the assignment are shown in Table S2. The matched structures in the database were shown, but the linkage and the 3- and 6-branch of the N-glycan mannose core were not specified.

glycan branching patterns in conjunction with the conventionally used MALDI-MS *N*-glycan profiling. When a *N*-glycan profile was obtained, detailed analysis by MS^n guided by the improved GIPS can be subsequently carried out for each of the *N*-glycans appeared in the mass profile as demonstrated by the profiling and analysis of *N*-glycans released from RNase B, adalimumab AGP and human serum. Lewis x epitope rather than core-fucosylation can be unambiguously identified. The method adds an additional dimension to *N*-glycan profiling by providing branching pattern analysis and thereafter the identification of individual *N*-glycans.

For reference glycan database, although there has been a considerable increase in number since 2000, e.g. the KEGG GLYCAN (Hashimoto et al., 2006), GlycomeDB (Ranzinger, Herget, von der Lieth, & Frank, 2011) and EUROCarbDB (von der Lieth et al., 2011), we selected the widely-used and well-documented CarbBank (Doubet & Albersheim, 1992) (also known as CCSD (Doubet, Bock, Smith, Darvill, & Albersheim, 1989, developed by the Complex Carbohydrate Research Center, University of Georgia Athens and consisted of 7837 glycan structures, for our use in this proof-of-concept work. Our approach presented here is not a *de novo* structural analysis tool. It relies very much on the repertoire of *N*-glycans presented in the database and can miss the glycan structure not included in the database. Although the development of CarbBank was discontinued in 1997 and this severely affected its use (Artemenko, McDonald, Davey, & Rudd, 2012), there have been international effort on further development (Tiemeyer et al., 2017), such as GlyTouCan (Aoki-Kinoshita et al., 2016) which is becoming the major open-access database for the glycoscience community. Unlike most of database searching approaches, the GIPS strategy is based on comparison glycan structures listed in the database rather than direct comparison of the mass spectra stored in the database.

In our present proof-of-concept work, only glycosidic cleavage is considered and used for assignment of branching patterns as the fragment ions thus produced contain little linkage information. There have been reports on the use of cross-ring fragmentation for partial linkage determination (Harvey, 2000; Harvey, Royle, Radcliffe, Rudd, & Dwek, 2008; Palma et al., 2015). Incorporation of these cross-ring fragmentations in future GIPS development will certainly help assignment of particular monosaccharides on specific *N*-glycan's antennae, and also improve the reliability and specificity of the method. The knowledges of the established biosynthetic pathways, glycosyltransferase activities, and the source of glycans are all important information to assist glycan assignment (Chen et al., 2016), and the linkage positions, e.g. the *N*-glycan 3- and 6-branches, may be specifically determined. These can be incorporated into the GIPS-II approach as specific parameters in future with potential for sequence and linkage analysis.

5. CRediT author statement

YL, SS, CH and WC conceived the study. SS, HW and DB designed the GIPS-II concept and computational model, and YL, CH and WC designed the mass spectrometry methodology. CH established the MALDI-MS^n and glycan structural analysis procedure and performed and analyzed the mass spectral data. HW, YW, JZ and JD implemented the GIPS-II approach. HG and JZ carried out glycan preparation and participated in MALDI-MS^n data acquisition. WC provided critical comments and advice on the GIPS-II concept and mass spectrometry experiments. CH, SS, WC, YL and DB wrote the manuscript. All authors discussed the results and commented on the manuscript.

Declaration of Competing Interest

The authors declare no competing financial interest.

Acknowledgment

This work was supported in part by the National Natural Science Foundation of China (Grant Nos. 31600650, 31671369 and 31770775), March of Dimes Prematurity Research Center grant (22-FY18-821), and Wellcome Trust Biomedical Resource grant (WT 218304/Z/19/Z).

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.carbpol.2020.116122.

Supplementary results in detail, three additional tables, seventeen figures about calculation process of GIPS-II approach for standards and glycoproteins are provided as supporting information

References

Aldredge, D., An, H. J., Tang, N., Waddell, K., & Lebrilla, C. B. (2012). Annotation of a serum N-Glycan library for rapid identification of structures. *Journal of Proteome Research, 11*, 1958–1968.

Anugraham, M., Jacob, F., Nixdorf, S., Everest-Dass, A. V., Heinzelmann-Schwarz, V., & Packer, N. H. (2014). Specific glycosylation of membrane proteins in epithelial ovarian cancer cell lines: Glycan structures reflect gene expression and DNA methylation status. *Molecular & Cellular Proteomics, 13*, 2213–2232.

Aoki-Kinoshita, K., Agravat, S., Aoki, N. P., Arpinar, S., Cummings, R. D., Fujita, A., et al. (2016). GlyTouCan 1.0-The international glycan structure repository. *Nucleic Acids Research, 44*, D1237–D1242.

Artemenko, A. V., McDonald, A. G., Davey, G. P., & Rudd, P. M. (2012). Databases and tools in glycobiology. *Methods in Molecular Biology (Clifton, NJ), 899*, 325–350.

Ashline, D. J., Singh, S., Hanneman, A., & Reinhold, V. (2005). Congruent strategies for carbohydrate sequencing. 1. mining structural details by MS^n. *Analytical Chemistry, 77*, 6250–6262.

Ashline, D. J., Lapadula, A. J., Liu, Y. H., Lin, M., Grace, M., Pramanik, B., et al. (2007). Carbohydrate structural isomers analyzed by sequential mass spectrometry. *Analytical Chemistry, 79*, 3830–3842.

Ashline, D. J., Zhang, H., & Reinhold, V. N. (2017). Isomeric complexity of glycosylation documented by MS^n. *Analytical and Bioanalytical Chemistry, 409*, 439–451.

Chai, W., Piskarev, V. E., Mulloy, B., Liu, Y., Evans, P. G., Osborn, H. M., et al. (2006). Analysis of chain and blood group type and branching pattern of sialylated oligosaccharides by negative ion electrospray tandem mass spectrometry. *Analytical Chemistry, 78*, 1581–1592.

Chen, Q., Pang, P. C., Cohen, M. E., Longtine, M. S., Schust, D. J., Haslam, S. M., et al. (2016). Evidence for differential glycosylation of trophoblast cell types. *Molecular & Cellular Proteomics, 15*, 1857–1866.

Dell, A., Reason, A. J., Khoo, K. H., Panico, M., McDowell, R. A., & Morris, H. R. (1994). Mass spectrometry of carbohydrate-containing biopolymers. *Methods in Enzymology, 230*, 108–132.

Doubet, S., & Albersheim, P. (1992). Letter to the glyco-forum. *Glycobiology, 2*, 505–507.

Doubet, S., Bock, K., Smith, D., Darvill, A., & Albersheim, P. (1989). The complex carbohydrate structure database. *Trends in Biochemical Sciences, 14*, 475–477.

Everest-Dass, A. V., Briggs, M. T., Kaur, G., Oehler, M. K., Homann, P., & Packer, N. H. (2016). N-glycan MALDI imaging mass spectrometry on formalin-fixed paraffin-embedded tissue enables the delineation of ovarian cancer tissues. *Molecular & Cellular Proteomics, 15*, 3003–3016.

Fournier, T., Medioubi-N, N., & Porquet, D. (2000). Alpha-1-acid glycoprotein. *Biochimica et Biophysica Acta, 1482*, 157–171.

Gao, C., Zhang, Y., Liu, Y., Feizi, T., & Chai, W. (2015). Negative-ion electrospray tandem mass spectrometry and microarray analyses of developmentally regulated antigens based on type 1 and type 2 backbone sequences. *Analytical Chemistry, 87*, 11871–11878.

Harvey, D. J. (2000). Electrospray mass spectrometry and fragmentation of N-linked carbohydrates derivatized at the reducing terminus. *Journal of the American Society for Mass Spectrometry, 11*, 900–915.

Harvey, D. J., Royle, L., Radcliffe, C. M., Rudd, P. M., & Dwek, R. A. (2008). Structural and quantitative analysis of N-linked glycans by matrix-assisted laser desorption ionization and negative ion nanospray mass spectrometry. *Analytical Biochemistry, 376*, 44–60.

Hashimoto, K., Goto, S., Kawano, S., Aoki-Kinoshita, K. F., Ueda, N., Hamajima, M., et al. (2006). KEGG as a glycome informatics resource. *Glycobiology, 16*, 63R–70R.

Hu, H., Khatrl, K., & Zaia, J. (2017). Algorithms and design strategies towards automated glycoproteomics analysis. *Mass Spectrometry Review, 36*, 475–498.

Huhn, C., Selman, M. H., Ruhaak, L. R., Deelder, A. M., & Wuhrer, M. (2009). IgG glycosylation analysis. *Proteomics, 9*, 882–913.

Joshi, H. J., Harrison, M. J., Schulz, B. L., Cooper, C. A., Packer, N. H., & Karlsson, N. G. (2004). Development of a mass fringerprinting tool for automated interpretation of oligosaccharide fragmentation data. *Proteomics, 4*, 1650–1664.

Ju, F., Zhang, J., Bu, D., Li, Y., Zhou, J., Wang, H., et al. (2019). De novo glycan structural identification from mass spectra using tree merging strategy. *Computational Biology and Chemistry, 80*, 217–224.

Kishino, S., Nomura, A., Itoh, S., Nakagawa, T., Takekuma, Y., Sugawara, M., et al.

(2002). Age- and gender-related differences in carbohydrate concentrations of α1-acid glycoprotein variants and the effects of glycoforms on their drug-binding capacities. *European Journal of Clinical Pharmacology, 58*, 621–628.
Kizuka, Y., & Taniguchi, N. (2016). Enzymes for *N*-glycan branching and their genetic and nongenetic regulation in cancer. *Biomolecule, 6*, 25.
Lohmann, K. K., & von der Lieth, C. W. (2004). Glycofragment and Glycosearchm-s: Web tools to support the interpretation of mass spectra of complex carbohydrates. *Nucleic Acids Research, 32*, W261 (Web Server issue).
Maass, K., Ranzinger, R., Geyer, H., von der Lieth, C.-W., & Geyer, R. (2007). "'Glycopeakfinder'" – De novo composition analysis of glycoconjugates. *Proteomics, 7*, 4435–4444.
Marth, J. D., & Grewal, P. K. (2008). Mammalian glycosylation in immunity. *Nature Reviews Immunology, 8*, 874–887.
Michael, C., & Rizzi, A. M. (2015). Quantitative isomer-specific N-glycan fingerprinting using isotope coded labeling and high performance liquid chromatography-electrospray ionization-mass spectrometry with graphitic carbon stationary phase. *Journal of Chromatography A, 1383*, 88–95.
Morimoto, K., Nishikaze, T., Yoshizawa, A. C., Kajihara, S., Aoshima, K., Oda, Y., et al. (2015). Glycan analysis plug-in: A database search tool for N-glycan structures using mass spectrometry. *Bioinformatics, 31*, 2217–2219.
Morris, T. A., Peterson, A. W., & Tarlov, M. J. (2009). Selective binding of RNase B glycoforms by polydopamine-immobilized concanavalin A. *Analytical Chemistry, 81*, 5413–5420.
Palma, A. S., Liu, Y., Zhang, Y., McCleary, B. V., Huang, Q., Guidolin, A. E., et al. (2015). Unravelling glucan recognition systems by glycome microarrays using the designer approach and mass spectrometry. *Molecular & Cellular Proteomics, 14*, 974–988.
Paulson, J. C., Blixt, O., & Collins, B. E. (2006). Sweet spots in functional glycomics. *Nature Chemical Biology, 2*, 238–248.
Pekelharing, J. M., Hepp, E., Kamerling, J. P., Gerwig, G. J., & Leijnse, B. (1988). Alterations in carbohydrate composition of serum IgG from patients with rheumatoid arthritis and from pregnant women. *Annals of the Rheumatic Diseases, 47*, 91–95.
Pompach, P., Brnakova, Z., Sanda, M., Wu, J., Edwards, N., & Goldman, R. (2013). Site-specific glycoforms of haptoglobin in liver cirrhosis and hepatocellular carcinoma. *Molecular & Cellular Proteomics, 12*, 1281–1293.
Raman, R., Raguram, S., Venkataraman, G., Paulson, J. C., & Sasisekharan, R. (2005). Glycomics: An integrated systems approach to structure-function relationships of glycans. *Nature Methods, 2*, 817–824.
Ranzinger, R., Herget, S., von der Lieth, C.-W., & Frank, M. (2011). GlycomeDB–a unified database for carbohydrate structures. *Nucleic Acids Research, 39*, D373–D376.
Reinhold, V. N., Zhang, H., Hanneman, A., & Ashline, D. (2013). Toward a platform for comprehensive glycan sequencing. *Molecular & Cellular Proteomics, 12*, 866–873.
Ren, S., Zhang, Z., Xu, C., Guo, L., Lu, R., Sun, Y., et al. (2016). Distribution of IgG galactosylation as a promising biomarker for cancer screening in multiple cancer types. *Cell Research, 26*, 963–966.
Reusch, D., Haberger, M., Maier, M., Kloseck, R., Zimmermann, B., Hook, M., et al. (2015). Comparison of methods for the analysis of therapeutic immunoglobulin G Fc-glycosylation profiles-part 1: separation-based methods. *mAbs, 7*, 167–179.
Rojas-Macias, M. A., Mariethoz, J., Andersson, P., Jin, C., Venkatakrishnan, V., Aoki, N. P., et al. (2019). Towards a standardized bioinformatics infrastructure for *N*- and *O*-glycomics. *Nature Communications, 10*, 3275.
Sarkar, A., Drouillard, S., Rivet, A., & Perez, S. (2015). Databases of conformations and NMR structures of glycan determinants. *Glycobiology, 25*, 1480–1490.
Sun, D., Hu, F., Gao, H., Song, Z., Xie, W., Wang, P., et al. (2019). Distribution of abnormal IgG glycosylation patterns from rheumatoid arthritis and osteoarthritis patients by MALDI-TOF-MS^{n}. *Analyst, 144*, 2042–2051.
Sun, S., Huang, C., Wang, Y., Liu, Y., Zhang, J., Zhou, J., et al. (2018). Toward automated identification of glycan branching patterns using multistage mass spectrometry with intelligent precursor selection. *Analytical Chemistry, 90*, 14412–14422.
Tiemeyer, M., Aoki, K., Paulson, J., Cummings, R. D., York, W. S., Karlsson, N. G., et al. (2017). GlyTouCan: An accessible glycan structure repository. *Glycobiology, 27*, 915–919.
Valk-Weeber, R. L., Dijkhuizen, L., & van Leeuwen, S. S. (2019). Large-scale quantitation of pure protein N-linked glycans. *Carbohydrate Research, 479*, 13–22.
von der Lieth, C. W., Freire, A. A., Blank, D., Campbell, M. P., Ceroni, A., Damerell, D. R., et al. (2011). EUROCarbDB: An open-access platform for glycoinformatics. *Glycobiology, 21*, 493–502.
Walsh, I., Zhao, S., Campbell, M., Taron, C.-H., & Rudd, P.-M. (2016). Quantitative profiling of glycans and glycopeptides: An informatics' perspective. *Current Opinion in Structural Biology, 40*, 70–80.
Wang, H., Zhang, J., Dong, J., Hou, M., Pan, W., Bu, D., et al. (2020). Identification of glycan branching patterns using multistage mass spectrometry with spectra tree analysis. *Journal of Proteomics, 217*, 103649.
Wang, Y., Bu, D., Huang, C., Wang, H., Zhou, J., Dong, J., et al. (2019). Best-first search guided multistage mass spectrometry-based glycan identification. *Bioinformatics, 35*, 2991–2997.
Wu, D., Struwe, W. B., Harvey, D. J., Ferguson, M. A. J., & Robinson, C. V. (2018). N-glycan microheterogeneity regulates interactions of plasma proteins. *Proceedings of the National Academy of Sciences of the United States of America, 115*, 8763–8768.
Yazawa, S., Takahashi, R., Yokobori, T., Sano, R., Mogi, A., Saniabadi, A. R., et al. (2016). Fucosylated glycans in α1-acid glycoprotein for monitoring treatment outcomes and prognosis of cancer patients. *PloS One, 11*, e0156277.
Zaia, J. (2008). Mass spectrometry and the emerging field of glycomics. *Chemistry & Biology, 15*, 881–892.