Mass spectrometry for glycan biomarker discovery

O. Aizpurua-Olaizola [a, b, *], J. Sastre Toraño [a], J.M. Falcon-Perez [b, c], C. Williams [b, c], N. Reichardt [d], G.-J. Boons [a]

[a] Department of Chemical Biology and Drug Discovery, Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Universiteitsweg 99, Utrecht, The Netherlands
[b] Exosomes Laboratory, CIC bioGUNE, CIBEREHD, Derio 48160 Bizkaia, Spain
[c] IKERBASKE Basque Foundation for Science, Bilbao, 48013 Bizkaia, Spain
[d] Glycotechnology Laboratory, CIC BiomaGUNE, Paseo Miramon 182, 20009 San Sebastian, Spain

ARTICLE INFO

ABSTRACT

The association between aberrant glycosylation of proteins and many cancers, and autoimmune and inflammatory diseases has been known for many years. Altered glycosylation can occur at the onset and during disease progression and identifying these changes at an early stage may greatly increase survival and improve quality of life. However, the identification of these biomarkers has not been easy, mainly due to the structural diversity and numerous possible glycan isomers. Fortunately, glycomics is becoming more feasible due to major improvements in mass spectrometry and separation science. The present review discusses recent methods for mass-spectrometry (MS) based glycomics for the identification of glycan biomarkers. Recent MS techniques with and without coupling to liquid chromatography, capillary electrophoresis or ion mobility spectrometry are described, and the most recent glycan biomarker studies are presented and future prospects discussed.

© 2017 Elsevier B.V. All rights reserved.

1. Relevance of glycans in biomarker discovery

Glycosylation is a common posttranslational protein modifications, and almost all cell surface and secreted proteins are modified by covalently-linked carbohydrates. Eukaryotic glycans are generally classified into two main groups: N- and O-glycans, where the glycan chains are linked to asparagine and serine/threonine residues, respectively (Fig. 1). N-glycans can be complex, hybrid or high-mannose type, and sharing a common pentasaccharide core, and presenting substantial heterogeneity due to variations in the number of antennae, terminal residues and core modifications. O-glycans consists of a larger variety of core structures, with a chain length ranging from 2 to 13 residues [1].
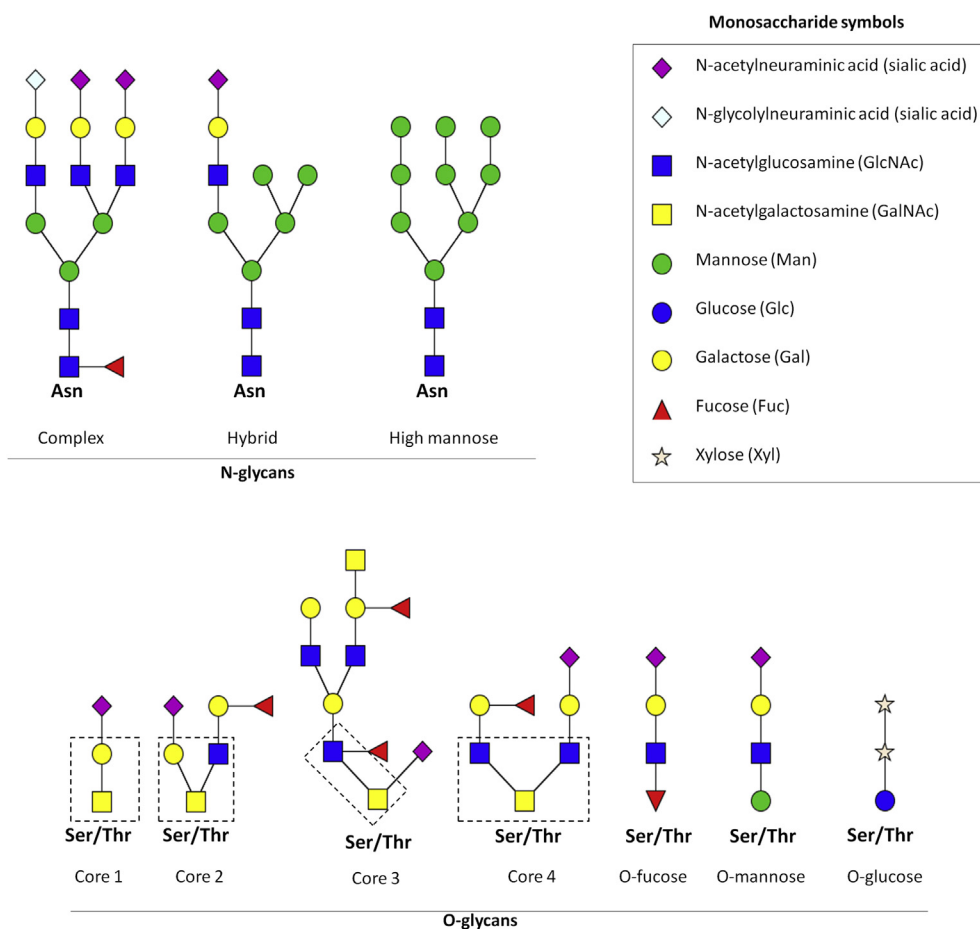
Glycans are essential mediators of biological processes such as protein folding, cell signalling, fertilization, embryogenesis, neuronal development, hormone activity and the proliferation of cells and their organization into specific tissues [2]. In addition, overwhelming data supports the relevance of glycosylation in pathogen recognition, inflammation, innate immune responses, and the development of autoimmune diseases and cancer [3]. Glycosylation is found on intracellular compartments, all cell surfaces and in extracellular matrix, which is composed of glycoproteins, proteoglycans, polysaccharides, and collagens. As a result glycans often create the first contact in cellular interactions [4]. Due to their structural diversity and functional roles, it is not surprising that some disease-related changes are more evident on glycan biosynthesis than that of protein.

Eukaryotic glycosylation is performed in the endoplasmic reticulum and *trans*-Golgi Network (TGN). It is a process that can

Fig. 1. Exemplary N- and O-glycan structures, which are linked to an asparagine residue and to a serine or threonine residue, respectively. Most common monosaccharide symbol nomenclature is represented.

reflect physiological state, and is dominated mostly by genetic factors and specific pathological situations, though also by other factors like age, gender and lifestyle [5]. In this sense, it is convenient to add an inter-subject variability study in any glycan biomarker study, as otherwise, factors like age and gender could lead to inaccurate conclusions. This can be fulfilled by defining subgroups according to these factors. On the other hand, due to a strong genetic control in glycosylation, intra-subject variability has been shown to be quite low in individuals over a short period of time [6]. This low intra-subject variability, along with inter-subject variability, makes glycans promising disease biomarkers for personalized medicine. Nevertheless, for the assessment of intra-subject variation longer studies are needed to provide a clearer picture, as sudden change of diet or habits could have a strong effect on glycan expression, complicating the identification of robust biomarkers. For a personalized treatment, patient stratification e.g. by age, gender, disease state, diet, genotype is required. This group will show a specific glycosylation profile that changes during beneficial treatment. In this respect, we are just beginning to unravel the functional significance of glycans in disease or treatment.

A glycan biomarker can be a single entity or be composed of a panel of compounds, either way it must be able to identify when the disease exists and when it does not, avoiding false positive and negative results. To make diagnostics easier to perform, biological samples should be obtained in a non-invasive manner, in particular for population screening. These samples could be biological fluids like plasma, serum or urine, that contain different glycosylated entities including cells, extracellular vesicles and proteins that can be interrogated after applying the corresponding isolation procedures. One of the most promising approaches in the field is to define specific glycans or set of glycans as cancer biomarkers, as several studies have shown that certain changes in plasma glycan structures occur during cancer initiation, progression, and treatment. A number of recent reviews describe the promise of glycans as cancer biomarkers [7–9]. An increase in fucosylation, sialylation, and branching of plasma N–glycans has been correlated with cancer development. In addition, it has been found that some glycans exhibit opposite trends in different cancers, as it is the case of hybrid glycans, with a lower expression in breast cancer and a higher expression in liver cancer. However, the utility of glycan profiles to determine organotropic metastasis has not been demonstrated yet.

Besides cancer biomarkers, another promising approach is the use of glycomics for detecting age-related diseases like Alzheimer's disease, Huntington's disease, progeroid syndromes, idiopathic normal pressure hydrocephalus, diabetes mellitus, frailty, and a variety of autoimmune/inflammatory diseases [10–12]. Glycans are involved in the pathophysiology of these diseases and, as a result, can be useful diagnostic biomarkers. For the discovery of such a specific patterns in an early disease state, it is necessary to aim for specific targets like specific cells or extracellular vesicles, as disease related changes in biological fluids would be too diluted to be detectable.

One of the major challenges in glycan biomarker discovery is the high inter-subject variability together with relatively low changes in the glycan profile in many cases. However, patient stratification could help overcoming this issue. On the other hand, cohorts and statistical analysis methods employed in studies differ considerably making it often difficult to draw general conclusions. Despite recent advances in glycomics, glycan identification is still a major bottle neck in developing non-invasive glycan biomarkers, due to the high analytical variability and the difficulty of determining exact structures of numerous possible glycan isomers that would require the generation of standards It is the expectation that with improved analytical approaches, and the availability of standards, the field of glycomics will have a greater potential for biomarker discovery for cancer and age-related diseases.

2. Current methods in glycan analysis for cancer biomarker discovery

All methods available for glycan analysis (reviewed in Ref. [13]) can in principle be applied to the biomarker discovery. Among these, MS-based techniques are a key for the recent expansion of glycomics capabilities. MS approaches, usually preceded by at least one separation technique, are necessary for more precise glycan characterizations. However, a complete structural assignment of unknown structures including sequence and linkage information is not possible by mass spectrometry unless standards are available for all possible isomeric structures. MS takes advantage of the underlying biochemical knowledge on well-defined glycan subsets like human N-glycosylation.

The analysis of released glycans usually involves several steps. First all glycans are released from glycoproteins; N-glycans by enzymatic release and O-glycans by chemical release. Specific exoglycosidases can also be used to obtain further information on linkage positions and stereochemistry. Next, the sample is purified usually by solid phase extraction (SPE) and glycans can be derivatized. The derivatization can be performed on the hydroxyl groups, reducing ends, and even on their associated sialic acids. In this way, charged or hydrophobic groups can be attached to glycans in order to enhance the separation and the MS detection. Moreover, derivatization processes such as permethylation can stabilize sialic acid residues, enhancing this way MS sensitivity, and helping in a detailed structural characterization by MS/MS. Different labels and their application were extensively reviewed by D.J. Harvey [14].

Before MS detection, and with the exception of matrix-assisted laser desorption/ionization (MALDI), glycans can be separated to reduce complexity and to avoid ion suppression. The most useful separation techniques include liquid chromatography (LC), mostly hydrophilic-interaction chromatography (HILIC) and porous graphitized carbon (PGC); capillary electrophoresis (CE) and ion mobility spectrometry (IMS), the latter usually coupled to either LC or CE.

High mass accuracy detectors, such as time-of-light (TOF), Orbitrap, or Fourier transform ion cyclotron resonance (FT-ICR)MS, can provide the data to annotate the compositions of the glycans, as low or even sub ppm mass errors are often required for exact mass assignment. However, many of the observed signals at particular m/z values cannot be assigned to a single positional isomer. As a result, tandem MS methods are commonly used to aid in the structural analyses. This tandem MS approaches can also be useful to avoid false compositional assignments due to high uncertainties associated with lower performance instruments.

Regarding MS data interpretation, reviews about databases and associated tools for glycomics were published recently [15,16]. A tool to obtain a monosaccharide composition is the online website GlycoMod, where compositions can be predicted from free or derivatized N- and O/glycans experimental MS data. On the other hand, the free software GlycoWorkbench can be used to predict theoretical MS/MS fragments of a selected structure. In addition, it provides a friendly interface for drawing structures and to search m/z values in databases. Another publicly available software is Glycolyzer, that provides tools for identifying glycan peaks in raw MS data, was demonstrated in ovarian cancer biomarker discovery [17]. In the following, we discuss recent MS approaches for glycan biomarker discovery.

2.1. MALDI-TOF MS

MALDI-TOF-MS is a commonly employed MS technique for N-glycan analysis, due to its high analysis speed, low sample consumption and high automation potential. It is often used as a first step for N-glycan profiling to obtain insight about possible structural diversity, or in large-scale studies where high speed and high throughput are demanded. Large-scale studies are necessary in biomarker studies especially when high inter subject variability can be an issue. However, MALDI-TOF-MS analysis does not involve a separation step, and therefore important information about less abundant species may be lost due to ion suppression, which will strongly depend on the matrix composition. In addition, it is less suitable for O-glycan analysis, as signals may overlap with the matrix background peaks in the lower mass range.

MALDI analyses results in dissociation of labile monosaccharide residues including sialic acids, and as a result a derivatization process is necessary, which can be challenging due to the propensity of sialic acid residues to form cyclic lactones. However, lactone formation can be used to differentiate α-2,3 from α-2,6 sialic acid residues, providing structural information otherwise not accessible. Permethylation is probably the most useful derivatization process, as it will enhance the ionization efficiency, stabilizes labile sialic acids, enables the detection of neutral and acidic glycans in positive mode, and facilitates the determination of branching and linkages positions. The downsides can be the incomplete derivatization, sample loss and considerably longer sample preparation times. This approach was used by Alley Jr. et al. to identify changes in the N-linked glycan structure expressions in the blood sera of ovarian cancer patients [18]. The total sera N-glycans and inmunoglobulin G (IgG)-associated N-glycans in samples from 20 healthy gender match controls and 19 women diagnosed with late-stage ovarian cancer were analysed. In addition, a more in-depth analysis of the fucosylation sites was carried out by exoglycosidase digestion. The authors observed an increase in the level of tri- and tetra-antennary structures, and galactosylated structures in cancer sera and IgG samples. Moreover, elevated levels of outer-arm fucosylation were also identified as opposed to changes in core fucosylation.

Derivatization requires laborious sample preparation and therefore is not practical for high-throughput studies, at least if it is not integrated into a conventional automated liquid handling platform. Therefore, an automated permethylation method for glycosylation analysis by MALDI-TOF-MS was developed recently [19]. Another promising derivatization technique is ethyl/methyl esterification, which can enable differentiation of 2,3 and 2,6-linked sialic acids [20]. This derivatization step was recently used by Gizaw et al. to study the N-glycan profile of Huntington's disease (HD) in transgenic mice [21]. They estimated the expression levels of 87 and 58 N-glycans in brain tissue and sera, respectively, of HD transgenic and control mice. Increased levels of core-fucosylated and bisecting-GlcNAc N-glycan types were found in the brain tissue of HD transgenic mice. Similarly, increased levels of

biantennary type glycans with core-fucosylated and sialic acid (particularly NeuGc) were found in HD serum samples. Differences in O-glycans were also noticed, with decreased core 1 type O-glycans and undetected core 2 type O-glycans for HD transgenic mice.

Glycans can also be labelled at the reducing end by reductive amination with 2-aminobenzoic acid (2-AA), 2-aminobenzamide acid (2-AB) or procainamide to allow negative mode MS detection [14]. This 2-AA labelling step was carried out by Balog et al. to study N-glycosylation in colorectal cancer tissue [22]. Structures with a bisecting GlcNAc were found to be less abundant in the tumours, whereas sulfated glycans, glycans with a sialylated Lewis type epitope, and paucimannosidic glycans were more abundant in tumour samples. In addition, core-fucosylated and high mannose N-glycans were identified in tumour tissues.

A different approach to avoid sample degradation during sample preparation is the on-target derivatization, which can be done using specific matrixes like 3-aminoquinoline [23]. Moreover, MALDI-FTICR can be used for glycan annotation to minimize in-source and metastable decay of sialic acids. This approach, along with 2-AA labelling, was recently used in a large-scale study of the total plasma N-glycome in order to find clinical markers of overall metabolic health and inflammation [24]. A number of 61 glycan compositions were relatively quantified in 2144 healthy middle-aged individuals, and associations with clinical variables were studied. Bisection, galactosylation, and sialylation of diantennary and tetraantennary species, and the size of high-mannose species were found to be important plasma markers associated with inflammation and overall metabolic health.

2.2. LC-ESI-MS

While being more time-consuming than MALDI-TOF glycan profiling, LC-ESI-MS analysis of released glycans is probably the most widely used approach to obtain comprehensive glycan profiles. LC techniques like reverse-phase (RP), Hydrophilic-interaction chromatography (HILIC), and Porous graphitized carbon (PGC) can be used for the separation of complex glycan mixtures. The availability of sub-2 μm particle size for stationary phases can alternatively provide increased resolution and sensitivity, along with shorter run times that can allow the implementation of LC-MS techniques in large-scale studies. Moreover, chip-based nano-LC−MS devices can greatly increase the reproducibility of glycomic analyses [25].

In addition to accurate MS measurements for compositional analysis, MS/MS fragmentation studies are very useful to obtain structural information of glycans separated by LC. A review covering the most important aspects of the characterization of isomeric glycan structures by LC-MS/MS was published recently [26].

For RP chromatography glycans must be derivatized with a hydrophobic label in order to have retention. The separation of glycans is highly dependent on the hydrophobicity of the label, but also on the contribution of the glycan due to different positions and linkages [27]. RP has the advantage that it is commonly used in almost all bioanalytical research laboratories, but the isomeric separation is limited in comparison to HILIC or PGC. Isomeric separation can be critical to find additional potential biomarkers, as showed by Huang et al. for distinguishing early-stage hepatocellular carcinoma from cirrhosis [28]. They permethylated released N-glycans from serum haptoglobin samples and made a comparison of RP (C18) and PGC columns. Both approaches enable the identification of several biomarkers, but with RP two specific isomers were masked by the other non-significant ones with the same m/z, and as a result these two biomarkers were missed.

The largest glycomics studies to date have been accomplished using HILIC chromatography [29]. A variety of HILIC columns can be used for LC-MS analysis. The most common stationary phases are silica-based ion exchange, zwitterionic (ZIC) and amide. Retention times are based on the hydrophilic potential of the molecule, which accounts for differences in properties that depend on size, charge, composition, structure, linkage and oligosaccharide branching. As a result, HILIC approaches are more useful for analysing the composition of complex mixtures, as many structural isomers can be separated prior to a tandem MS analysis [30]. When using HILIC, labelling of the reducing end by reductive amination is often the preferred derivatization approach because it reduces complexity at the reducing anomeric center. This approach was employed by Giménez et al. to develop a methodology for cancer biomarker discovery [31]. They carried out isotopic labelling by reductive amination using [12C]- and [13C]-coded aniline for control samples and pathological samples, respectively. Then, a control sample was pooled to each pathological sample in order to compare sample populations without the influence of ion suppression effects and the tuning of the mass spectrometer. μZIC-HILIC-ESI-MS technique was used for the separation and analysis. The methodology was applied to distinguish between inflammatory diseases of the pancreas, pancreatic cancer, and healthy controls.

PGC behaves as a stronger reversed-phase than the traditional alkyl-bonded silica. The retention mechanism of carbohydrates is mainly based on adsorption and the flat surface of PGC packings brings about unique selectivity, but also includes hydrophobic interactions. The elution patterns are based on the size and planarity of the molecule (position and configuration of linkage) which makes this technique appropriate for the separation of complex and positional isomeric glycans. Similarly to RP, hydrophobic derivatizations like permethylation can be performed. Recently, a review on clinical glycomics employing PGC-LC-MS was published [32]. This separation technique was used by Sethi et al. to identify N-glycan biomarkers in colorectal cancer (CRC) [33]. They analysed cell line- and tissue-derived native N-glycans, reduced with sodium borohydride to convert α and β anomers of the reducing end to sugar alditols. They used accurate mass measurements and the Glycomod tool to identify glycan compositions, and MS/MS data generated by collisional induced dissociation (CID) and the Glyco-Workbench software for tentative structural assignment. An over-expression of high mannose and α2,6-sialylated glycans in CRC was observed. Moreover, specific N-glycan markers for different metastatic phenotypes, and tumour staging were found.

A similar approach was performed to study the glycosylation patterns of extreme longevity and healthy aging [34]. Using a PGC-based nano-LC system coupled to an FT-ICR ion trap MS system, plasma derived reduced N-glycan profiles were established for 6 centenarian females and 10 control aged and young females. The results showed that higher-branched and highly sialylated N-glycans as well as agalacto- and/or bisecting N-glycans were increased in centenarians, while biantennary N-glycans were decreased. These changes were attributed to enhanced chronic inflammations observed in centenarians, suggesting that responses to inflammation may play an important role in extreme longevity and healthy aging in humans.

In summary, it must be highlighted that PGC exhibits an enhanced isomeric separation compared to other separation techniques like HILIC or RP. Moreover, the chemical background noise is lower with PGC than with HILIC, due to the higher content of water in the mobile phase. However, application of PGC to large scale glycomics studies is somehow limited by robustness problems, as polar compounds can be accumulated over time, resulting in retention time drift [35].

2.3. CE-MS

Capillary gel electrophoresis (CGE) coupled to laser-induced fluorescence (CGE-LIF) had allowed for high sensitive glycomics since the early 1990s [36]. Indeed, despite improvements in instrumentation and methods based on MS, CGE-LIF is still the key CE technique for glycan profiling. This is due to the extraordinary detection limits offered by LIF detection and the high resolution and migration time repeatability obtained with CGE, allowing for the separation of many structural isomers and structure assignment through the application of glucose unit ladders [37]. There is still a need for more detailed structural information that could be obtained with MS, but the gels nowadays applied in CGE are incompatible with MS, resulting in little developments in the CGE-MS area. However, during the past few years, several advances in CE-MS have been made on "MS friendly" buffer systems, pre-concentration methods, and CE-MS interfacing designs like the sheathless porous tip interface and the flow-through microvial interface [38].

CE-TOF-MS using a porous tip sheathless interface was used to study the glycoform profiles of pharmaceutical proteins, obtaining a similar coverage to sheath–liquid CE-MS while injecting ten times less sample [39]. On the other hand, flow-through microvial interface was evaluated for native neutral and sialylated N-glycans from human erythropoietin (rHuEPO) using a triple quadrupole MS and an ultra-high resolution TOF-MS [40]. The method was compared to a well-established PGC-LC-MS/MS method, and more features were identified by CE than with PGC, even though the separation of native glycans is less efficient than glycans labelled with 8-aminopyrene-1,3,6-trisulfonic acid (APTS). The major limitation of the method is probably the limited concentration sensitivity, and therefore concentrated samples for a complete glycan profile are necessary. In order to improve the low concentration sensitivity, an in-line SPE-CE-MS method using a fritless bead string design for APTS labelled glycans can be used [41]. In this way enrichment factor of more than 800 could be achieved compared to the corresponding CE-MS method without preconcentration.

Regarding "MS friendly" systems, Bunz et al. studied various neutral coated capillaries, buffer system, and co-axial sheath liquid compositions for CE-TOF-MS analysis of N-glycans [42]. Released N-glycans were labelled by reductive amination with APTS to introduce a uniform negative charge required for electrophoretic mobility. The separation efficiency and the sensitivity of the optimized method were found similar to routine CE-LIF systems.

Khatri et al. performed the analysis of released glycans on a microfluidic CE-ESI device, with integrated sprayer, using aminoxy Tandem Mass Tag (TMT) and sialic acid derivatization to improve separation and enable multiplexed quantitation by MS/MS [43]. It is clear that the lack of standards makes it necessary to develop alternatives for quantification and therefore Váradi et al. developed a CE-TOF-MS method for the analysis of N-glycans labelled with $^{12}C_6$ and $^{13}C_6$ 2-AA isotopes [44]. This approach offers the possibility to perform accurate comparability studies between populations due to minimized technical variation with isotope labelling.

The ability to resolve structural isomers and the minute amount of sample needed makes CE-MS a powerful method for glycan biomarker discovery, alongside or even orthogonally to LC-MS or MALDI-MS. Although the application of CE-MS for glycan biomarker discovery has been limited, with the advances in "MS friendly" buffer systems, new preconcentration methods, and innovative interfacing technologies it is expected that it will play a key role in the near future.

2.4. IMS-MS

IMS-MS is a relatively new technique that can enhance the structural characterization of glycans due to its isomeric separation potential [45–47]. IMS measures the mobility of ions in the gas-phase under the influence of an electric field, which is based on the shape, size and charge of the ions. Measurements are performed on a milliseconds timescale. MS/MS experiments of separated isomers can provide structural information, like the identification of sialic acid linkage characteristics (α2-3 or α2-6) [48], or the identification of epimeric glycans [49]. Moreover, rotationally averaged collision cross section (CCS) values can be obtained. CCS values are important distinguishing characteristics of ions in the gas phase, and in addition to the empirical determinations it can also be calculated computationally when the 3D structure of the molecule is known. This feature makes IMS a very promising tool, as adding CCS values of glycans and their fragments to databases will increase structural identification confidence and accuracy. Several approaches have been made recently to build released glycans CCS databases [50–52]. However, in order to build a database that can allow a complete structural assignment of unknown glycans, synthetic standards with known structures and linkage configurations will be necessary. In addition to provide CCS values, IMS can be coupled to other separation techniques such as LC and CE, providing orthogonal separation (unpublished data) (Fig. 2).

Several types of ion mobility analyzers have been coupled to MS, such as drift tube (DT)IMS, travelling wave (TW)IMS, and trapped (T)IMS. These three analyzers operate in low field conditions, so unlike other analyzers operating in high field conditions, CCS values can be derived from the mobility data. In DTIMS ions are introduced into a cell filled with a static drift gas, in which a weak electric field is applied causing the ion packages to move through the cell. Ions that undergo more interactions with the drift gas have higher drift times and as a result higher CCS values. A custom-made DTIMS-TOFMS instrument was successfully applied to create N-glycan mobility profiles that were used to distinguish between Barrett's oesophagus disease, high-grade dysplasia, oesophageal adenocarcinoma, and normal control samples [53]; as well as between liver cirrhosis, liver cancer, and control samples [54]. There are few glycan analysis studies with DTIMS, but the number is expected to rise after the recent release of the first commercial instrument, the Agilent 6560 Ion Mobility Q-TOF (Fig. 3). CCS values are automatically calibrated using a tune mix with standard CCS reference values, and the instrument can be coupled to LC and or CE. This equipment has been already used for N-glycan analysis, coupled to the Agilent HPLC-chip instrument [52].

The first TWIMS instrument was released by Waters in 2006 (Fig. 3). In this technique ions are directed across a cell filled with a static gas using series of ring electrodes. The travelling waves are generated by increasing the voltage on periodically spaced electrodes and then transferring the potential along the ring electrodes. An additional RF potential is applied to confine the ions. Ions with lower interaction with the drift times will "surf" the waves, and as a result have lower CCS values. These CCS values are obtained by manual calibration against molecules whose drift tube CCS values are known, and if possible of a similar structure to the analytes being studied. In recent studies, dextran was used as a calibrant to calculate CCS values for negatively [50] and positively [55] charged N-glycans and their fragments.

The first commercial instrument based on TIMS technology has been recently released by Bruker (Fig. 3). In this instrument, an axial electric field and an opposite drift gas flow is applied to trap
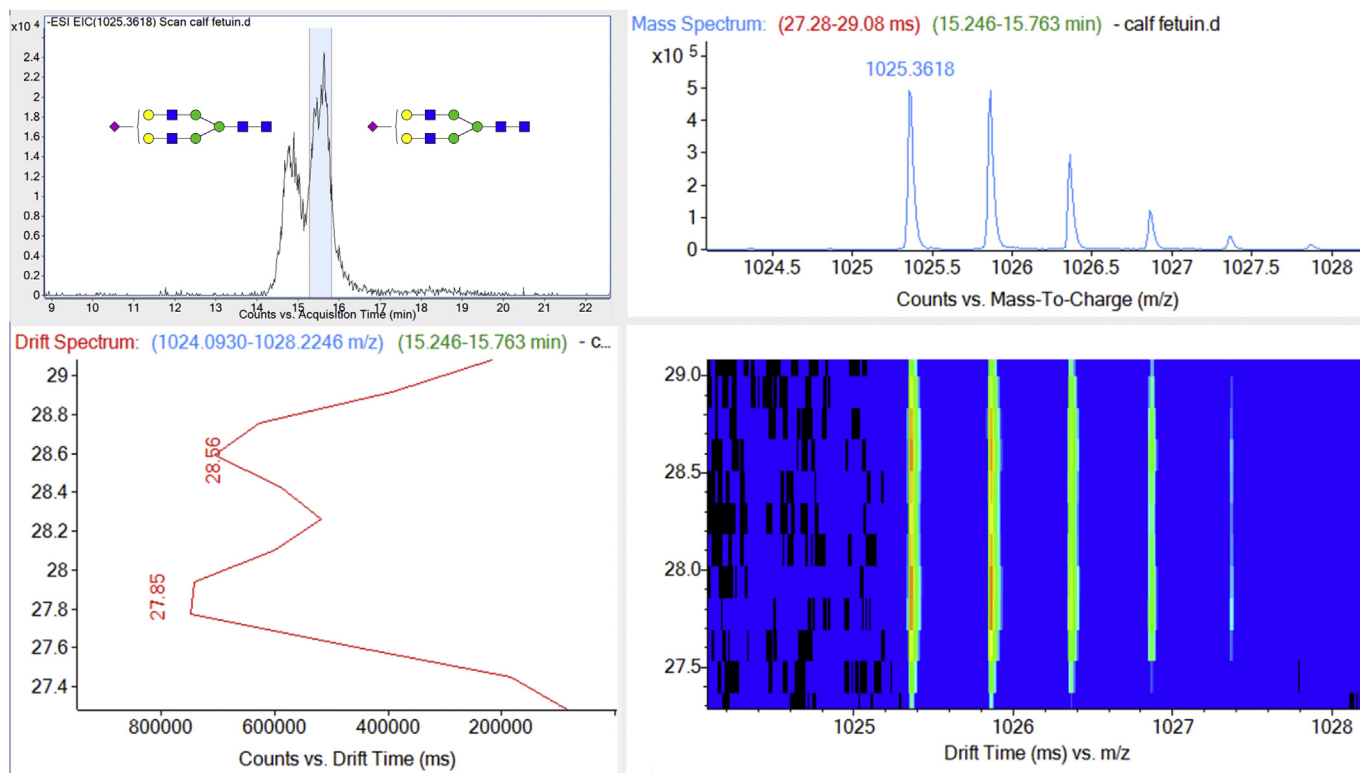
Fig. 2. 2D plot of the drift time (C) against *m/z* (D) of the 2-AA labelled negatively charged ions from calf fetuin, analysed by LC-IMS-TOF MS using a ZIC-HILIC column. A selected range of the chromatogram is shown in (A), and mass spectrum in (B). 2 isomers are clearly separated by HILIC (A), whereas IMS provides extra features (C and D) that HILIC cannot achieve, as it is shown for the selected peak in (A). Unpublished data from Aizpurua-Olaizola et al.

ions at different regions of the cell depending the ions CCS values. At the same time, a quadrupole field perpendicular to the electric field is created by RF potential. The axial electric field is then gradually lowered eluting ions from high CCS values to low CCS values. Similarly to TWIMS, these values are obtained by calibration with known standards. This methodology was used by Pu et al. to separate and identify glycan linkage isomers by performing tandem MS analysis by FTICR on mobility-selected ions for confident analyte identification [56].

3. The future of glycan biomarker discovery

MS techniques have turned into indispensable tools for glycan biomarker discovery, due to its high accurate mass determination for compositional analysis, the ability of MS/MS fragmentation techniques for structural identification, and the possibility to couple with separation techniques like LC, CE, and IMS. These integrated methods have great potential for isomeric separation, which may be essential to find novel biomarkers. Moreover, the possibility to couple multidimensionally LC or CE with IMS can provide extra isomeric separation, along with CCS values that can be implemented into existing glycan databases. Recently, a promising methodology to resolve carbohydrate isomers has also been published combining MS and infrared (IR) spectroscopy [57]. The authors highlighted the complementarity of IR spectroscopy with IMS and qualify the IR fingerprint in the 3 μm spectral range as a powerful metric for structural characterization of carbohydrates. Besides, improvements in automation, from sample pretreatment to data analysis, open the possibility to use these techniques in large-scale studies and in clinical practice.

Still, data analysis is the main bottleneck in biomarker discovery studies, as obtaining a complete structural elucidation of complex glycan mixtures and achieving quantitative measurements are still real challenges. It is expected that synthetic pure glycan standards will play a major role in the near future, both, for structural identification, but also for quantitation using isotopically labelled standards [58]. Glycan composition not only changes qualitatively, but also in absolute terms, allowing the possibility to implement clinical cut-off values for several diseases [59]. Some labelled standards are already available for routine analysis, mostly for robust quantitative clinical methods than for complex samples commonly used in biomarker discovery. In this sense, the recently developed strategy for the chemoenzymatic synthesis of complex branched N-glycans can provide the needed standards in the near future [60,61]. Moreover, algorithms for glycan structure assignment based on fragmentation, CCS and chromatographic properties must be developed. This will allow the identification of exact structures of glycans in biological samples without the need of a large number of synthetic standards.

When biomarker discovery studies focus on total serum or on a total tissue differences rather than on protein-specific differences, potential biomarkers from specific proteins may be diluted. It is expected that a separate glycan profiling of promising key proteins will be the pathway to discover novel biomarkers. Recently, it was found that integrins located in the membranes of extracellular vesicles could be used to predict organ-specific metastasis [62]. In this sense, glycoprofiling of these extracellular vesicles and/or integrins could also be a promising approach for organ-specific cancer biomarker discovery.
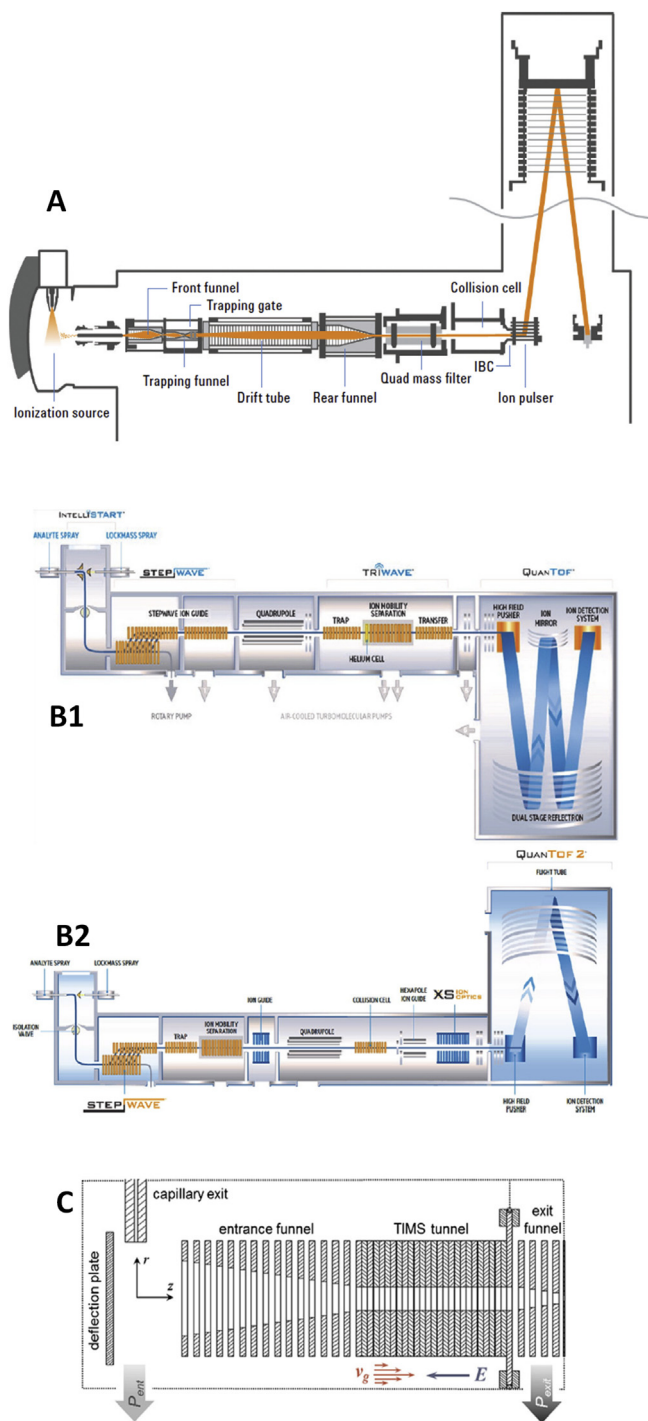
Fig. 3. Schematics of the instruments Agilent 6560 Ion Mobility Q-TOF (A), Waters Synapt G2-Si (B1) and Vion IMS Q-TOF (B2), and Bruker timsTOF (C).

Acknowledgements

References

[1] C. Jin, et al., Structural diversity of human gastric mucin glycans, Mol. Cell. Proteom. (2017). https://doi.org/10.1074/mcp.M117.067983.

[2] M.A. Wolfert, G.J. Boons, Adaptive immune activation: glycosylation does matter, Nat. Chem. Biol. 9 (2013) 776—784. https://doi.org/10.1038/nchembio.1403.

[3] M. Dalziel, M. Crispin, C.N. Scanlan, N. Zitzmann, R.A. Dwek, Emerging principles for the therapeutic exploitation of glycosylation, Science 343 (2014), 1235681. https://doi.org/10.1126/science.1235681.

[4] J.Q. Gerlach, M.D. Griffin, Getting to know the extracellular vesicle glycome, Mol. Biosyst. 4 (2016) 1071—1081. https://doi.org/10.1039/c5mb00835b.

[5] A. Knezevic, et al., Effects of aging, body mass index, plasma lipid profiles, and smoking on human plasma N-glycans, Glycobiology 20 (2010) 959—969. https://doi.org/10.1093/glycob/cwq051.

[6] O. Gornik, J. Wagner, M. Pucić, A. Knezević, I. Redzic, G. Lauc, Stability of N-glycan profiles in human plasma, Glycobiology 19 (2009) 1547—1553. https://doi.org/10.1093/glycob/cwp134.

[7] M.J. Kailemia, D. Park, C.B. Lebrilla, Glycans and glycoproteins as specific biomarkers for cancer, Anal. Bioanal. Chem. 409 (2017) 395—410. https://doi.org/10.1007/s00216-016-9880-6.

[8] Y. Lan, C. Hao, X. Zeng, Y. He, P. Zeng, Z. Guo, L. Zhang, Serum glycoprotein-derived N- and O-linked glycans as cancer biomarkers, Am. J. Cancer Res. 6 (2016) 2390—2415.

[9] N. Taniguchi, Y. Kizuka, Glycans and cancer: role of N-glycans in cancer biomarker, progression and metastasis, and therapeutics, Adv. Cancer Res. 126 (2015) 11—51. https://doi.org/10.1016/bs.acr.2014.11.001.

[10] Y. Miura, T. Endo, Glycomics and glycoproteomics focused on aging and age-related diseases—Glycans as a potential biomarker for physiological alterations, Biochim. Biophys. Acta 1860 (2016) 1608—1614. https://doi.org/10.1016/j.bbagen.2016.01.013.

[11] E. Maverakis, et al., Glycans in the immune system and the altered glycan theory of autoimmunity: a critical review, J. Autoimmun. 57 (2015) 1—13. https://doi.org/10.1016/j.jaut.2014.12.002.

[12] F. Dall'Olio, V. Vanhooren, C.C. Chen, P.E. Slagboom, M. Wuhrer, C. Franceschi, N-glycomic biomarkers of biological aging and longevity: a link with inflammaging, Ageing Res. Rev. 12 (2013) 685—698. https://doi.org/10.1016/j.arr.2012.02.002.

[13] W.R. Alley Jr., B.F. Mann, M.V. Novotny, High-sensitivity analytical approaches for the structural characterization of glycoproteins, Chem. Rev. 113 (2013) 2668—2732. https://doi.org/10.1021/cr3003714.

[14] D.J. Harvey, Derivatization of carbohydrates for analysis by chromatography; electrophoresis and mass spectrometry, J. Chromatogr. B 879 (2011) 1196—1225. https://doi.org/10.1016/j.jchromb.2010.11.010.

[15] F. Lisacek, et al., Databases and associated tools for glycomics and glyco-proteomics, Methods Mol. Biol. 1503 (2017) 235—264. https://doi.org/10.1007/978-1-4939-6493-2_18.

[16] I. Walsh, S. Zhao, M. Campbell, C.H. Taron, P.M. Rudd, Quantitative profiling of glycans and glycopeptides: an informatics' perspective, Curr. Opin. Struct. Biol. 40 (2016) 70—80. https://doi.org/10.1016/j.sbi.2016.07.022.

[17] S.R. Kronewitter, et al., The glycolyzer: automated glycan annotation software for high performance mass spectrometry and its application to ovarian cancer glycan biomarker discovery, Proteomics 12 (2012) 2523—2538. https://doi.org/10.1002/pmic.201100273.

[18] W.R. Alley Jr., et al., N-linked glycan structures and their expressions change in the blood sera of ovarian cancer patients, J. Proteome Res. 11 (2012) 2282—2300. https://doi.org/10.1021/pr201070k.

[19] A. Shubhakar, A. Shubhakar, R.P. Kozak, K.R. Reiding, L. Royle, D.I. Spencer, D.L. Fernandes, M. Wuhrer, Automated high-throughput permethylation for glycosylation analysis of biologics using MALDI-TOF-MS, Anal. Chem. 88 (2016) 8562—8569. https://doi.org/10.1021/acs.analchem.6b01639.

[20] K.R. Reiding, E. Lonardi, A.L. Hipgrave Ederveen, M. Wuhrer, Ethyl esterification for MALDI-MS analysis of protein glycosylation, Methods Mol. Biol. 1394 (2016) 151—162. https://doi.org/10.1007/978-1-4939-3341-9_11.

[21] S.T. Gizaw, T. Koda, M. Amano, K. Kamimura, T. Ohashi, H. Hinou, S. Nishimura, A comprehensive glycome profiling of Huntington's disease transgenic mice, Biochim. Biophys. Acta 1850 (2015) 1704—1718. https://doi.org/10.1016/j.bbagen.2015.04.006.

[22] C.I. Balog, et al., N-glycosylation of colorectal cancer tissues, Mol. Cell. Proteom. 11 (2012) 571—585. https://doi.org/10.1074/mcp.M111.011601.

[23] T. Nishikaze, K. Kaneshiro, S. Kawabata, K. Tanaka, Structural analysis of N-glycans by the glycan-labeling method using 3-aminoquinoline-based liquid matrix in negative-ion MALDI-MS, Anal. Chem. 84 (2012) 9453—9461. https://doi.org/10.1021/ac302286e.

[24] K.R. Reiding, et al., Human plasma N-glycosylation as analyzed by matrix-assisted laser desorption/ionization-fourier transform ion cyclotron resonance-MS associates with markers of inflammation and metabolic health, Mol. Cell. Proteom. 16 (2017) 228—242. https://doi.org/10.1074/mcp.M116.065250.

[25] S. Hua, C. Lebrilla, H.J. An, Application of nano-LC-based glycomics towards biomarker discovery, Bioanalysis 3 (2011) 2573—2585. https://doi.org/10.4155/bio.11.263.

[26] L. Veillon, Y. Huang, W. Peng, X. Dong, B.G. Cho, Y. Mechref, Characterization of isomeric glycan structures by LC-MS/MS, Electrophoresis (2017). https://doi.org/10.1002/elps.201700042.

[27] G.C.M. Vreeker, M. Wuhrer, Reversed-phase separation methods for glycan analysis, Anal. Bioanal. Chem. 409 (2017) 359—378. https://doi.org/10.1007/s00216-016-0073-0.

[28] Y. Huang, S. Zhou, J. Zhu, D.M. Lubman, Y. Mechref, LC-MS/MS isomeric profiling of permethylated N-glycans derived from serum haptoglobin of hepatocellular carcinoma (HCC) and cirrhotic patients, Electrophoresis (2017). https://doi.org/10.1002/elps.201700025.

[29] C. Lauc, et al., Genomics meets glycomics-the first GWAS study of human N-Glycome identifies HNF1α as a master regulator of plasma protein fucosylation, PLoS Genet. (2010), e1001256. https://doi.org/10.1371/journal.pgen.1001256.

[30] G. Zauner, A.M. Deelder, M. Wuhrer, Recent advances in hydrophilic interaction liquid chromatography (HILIC) for structural glycomics, Electrophoresis 32 (2015) 3456—3466. https://doi.org/10.1002/elps.201100247.

[31] E. Giménez, et al., Quantitative analysis of N-glycans from human alfa-acid-glycoprotein using stable isotope labeling and zwitterionic hydrophilic interaction capillary liquid chromatography electrospray mass spectrometry as tool for pancreatic disease diagnosis, Anal. Chim. Acta 2015 (866) (2015) 59—68. https://doi.org/10.1016/j.aca.2015.02.008.

[32] K. Stavenhagen, D. Kolarich, M. Wuhrer, Clinical glycomics employing graphitized carbon liquid chromatography-mass spectrometry, Chromatographia 78 (2015) 307—320. https://doi.org/10.1007/s10337-014-2813-7.

[33] M.K. Sethi, W.S.2 Hancock, S. Fanayan, Identifying N-Glycan biomarkers in colorectal cancer by mass spectrometry, Acc. Chem. Res. 49 (2016) 2099—2106. https://doi.org/10.1021/acs.accounts.6b00193.

[34] Y. Miura, et al., Change in N-Glycosylation of plasma proteins in Japanese semisupercentenarians, PLoS One 10 (2015), e0142645. https://doi.org/10.1371/journal.pone.0142645.

[35] L. Mauko, N.A. Lacher, M. Pelzing, A. Nordborg, P.R. Haddad, E.F. Hilder, Comparison of ZIC-HILIC and graphitized carbon-based analytical approaches combined with exoglycosidase digestions for analysis of glycans from monoclonal antibodies, J. Chromatogr. B 911 (2012) 93—104. https://doi.org/10.1016/j.jchromb.2012.10.043.

[36] C. Schwedler, et al., Identification of 34 N-glycan isomers in human serum by capillary electrophoresis coupled with laser-induced fluorescence allows improving glycan biomarker discovery, Anal. Bioanal. Chem. 406 (2014) 7185—7193. https://doi.org/10.1007/s00216-014-8168-y.

[37] M. Kerékgyártó, A. Guttman, Toward the generation of an aminonaphthalene trisulfonate labeled N-glycan database for capillary gel electrophoresis analysis of carbohydrates, Electrophoresis 35 (2014) 2222—2228. https://doi.org/10.1002/elps.201400054.

[38] P.W. Lindenburg, R. Haselberg, G. Rozing, R. Ramautar, Developments in interfacing designs for CE—MS: towards enabling tools for proteomics and metabolomics, Chromatographia 78 (2015) 367—377. https://doi.org/10.1007/s10337-014-2795-5.

[39] R. Haselberg, G.J. de Jong, G.W. Somsen, Low-flow sheathless capillary Electrophoresis—Mass spectrometry for sensitive glycoform profiling of intact pharmaceutical proteins, Anal. Chem. 85 (2013) 2289—2296. https://doi.org/10.1021/ac303158f.

[40] R.G. Jayo, et al., Simple capillary electrophoresis—mass spectrometry method for complex glycan analysis using a flow-through microvial interface, Anal. Chem. 86 (2014) 6479—6486. https://doi.org/10.1021/ac5010212.

[41] K. Jooß, J. Sommer, S.C. Bunz, C. Neusüß, In-line SPE-CE using a fritless bead string design—application for the analysis of organic sulfonates including inline SPE-CE-MS for APTS-labeled glycans, Electrophoresis 35 (2014) 1236—1243. https://doi.org/10.1002/elps.201300388.

[42] S.C. Bunz, E. Rapp, C. Neusüss, Capillary electrophoresis/mass spectrometry of APTS-labeled glycans for the identification of unknown glycan species in capillary electrophoresis/laser-induced fluorescence systems, Anal. Chem. 85 (2013) 10218—10224. https://doi.org/10.1021/ac401930j.

[43] K. Khatri, J.A. Klein, J.R. Haserick, D.R. Leon, C.E. Costello, M.E. McComb, J. Zaia, Microfluidic capillary electrophoresis-mass spectrometry for analysis of monosaccharides, oligosaccharides, and glycopeptides, Anal. Chem. 89 (2017) 6645—6655. https://doi.org/10.1021/acs.analchem.7b00875.

[44] C. Váradi, S. Mittermayr, S. Millán-Martín, J. Bones, Quantitative twoplex glycan analysis using $^{12}C_6$ and $^{13}C_6$ stable isotope 2-aminobenzoic acid labelling and capillary electrophoresis mass spectrometry, Anal. Bioanal. Chem. 408 (2016) 8691—8700. https://doi.org/10.1007/s00216-016-9935-8.

[45] J. Hofmann, K. Pagel, Glycan analysis by ion mobility-mass spectrometry, Angew Chem. Int. Ed Engl. 56 (2017) 8342—8349. https://doi.org/10.1002/anie.201701309.

[46] J. Hofmann, H.S. Hahm, P.H. Seeberger, K. Pagel, Identification of carbohydrate anomers using ion mobility-mass spectrometry, Nature 526 (2015) 241—244. https://doi.org/10.1038/nature15388.

[47] C.J. Gray, B. Thomas, R. Upton, L.G. Migas, C.E. Eyers, P.E. Barran, S.L. Flitsch, Applications of ion mobility mass spectrometry for high throughput, high resolution glycan analysis, Biochim. Biophys. Acta 1860 (2016) 1688—1709. https://doi.org/10.1016/j.bbagen.2016.02.003.

[48] M. Guttman, K.K. Lee, Site-specific mapping of sialic acid linkage isomers by ion mobility spectrometry, Anal. Chem. 88 (2017) 5212—5217. https://doi.org/10.1021/acs.analchem.6b00265.

[49] P. Both, et al., Discrimination of epimeric glycans and glycopeptides using IM-MS and its potential for carbohydrate sequencing, Nat. Chem. 6 (2014) 65—74. https://doi.org/10.1038/nchem.1817.

[50] J. Hofmann, W.B. Struwe, C.A. Scarff, J.H. Scrivens, D.J. Harvey, K. Pagel, Estimating collision cross sections of negatively charged N-glycans using traveling wave ion mobility-mass spectrometry, Anal. Chem. 86 (2014) 10789—10795. https://doi.org/10.1021/ac5028353.

[51] W.B. Struwe, K. Pagel, J.L. Benesch, D.J. Harvey, M.P. Campbell, GlycoMob: an ion mobility-mass spectrometry collision cross section database for glycomics, Glycoconj. J. 33 (2016) 399—404. https://doi.org/10.1007/s10719-015-9613-7.

[52] R.S. Glaskin, K. Khatri, Q. Wang, J. Zaia, C.E. Costello, Construction of a database of collision cross section values for glycopeptides, glycans, and peptides determined by IM-MS, Anal. Chem. 89 (2017) 4452—4460. https://doi.org/10.1021/acs.analchem.6b04146.

[53] M.M. Gaye, S.J. Valentine, Y. Hu, N. Mirjankar, Z.T. Hammoud, Y. Mechref, B.K. Lavine, D.E. Clemmer, Ion mobility-mass spectrometry analysis of serum N-linked glycans from esophageal adenocarcinoma phenotypes, J. Proteome Res. 11 (2012) 6102—6110. https://doi.org/10.1021/pr300756e.

[54] D. Isailovic, et al., Delineating diseases by IMS-MS profiling of serum N-linked glycans, J. Proteome Res. 11 (2012) 576—585. https://doi.org/10.1021/pr200777u.

[55] J. Hofmann, A. Stuckmann, M. Crispin, D.J. Harvey, K. Pagel, W.B. Struwe, Identification of Lewis and blood group carbohydrate epitopes by ion mobility-tandem-mass spectrometry fingerprinting, Anal. Chem. 89 (2017) 2318—2325. https://doi.org/10.1021/acs.analchem.6b03853.

[56] Y. Pu, M.E. Ridgeway, R.S. Glaskin, M.A. Park, C.E. Costello, C. Lin, Separation and identification of isomeric glycans by selected accumulation-trapped ion mobility spectrometry-electron activated dissociation tandem mass spectrometry, Anal. Chem. 88 (2016) 3440—3443. https://doi.org/10.1021/acs.analchem.6b00041.

[57] B. Schindler, et al., Anomeric memory of the glycosidic bond upon fragmentation and its consequences for carbohydrate sequencing, Nat. Commun. 8 (2017) 973. https://doi.org/10.1038/s41467-017-01179-y.

[58] J. Etxebarria, N.C. Reichardt, Methods for the absolute quantification of N-glycan biomarkers, Biochim. Biophys. Acta 1860 (2017) 1676—1687. https://doi.org/10.1016/j.bbagen.2016.03.003.

[59] Y. Ishibashi, et al., Serum tri- and tetra-antennary N-glycan is a potential predictive biomarker for castration-resistant prostate cancer, Prostate 74 (2014) 1521—1529. https://doi.org/10.1002/pros.22869.

[60] Z. Wang, Z.S. Chinoy, S.G. Ambre, W. Peng, R. McBride, R.P. de Vries, J. Glushka, J.C. Paulson, G.J. Boons, A general strategy for the chemoenzymatic synthesis of asymmetrically branched N-glycans, Science 341 (2013) 379—383. https://doi.org/10.1126/science.1236231.

[61] I.A. Gagarinov, T. Li, J.S. Toraño, T. Caval, A.D. Srivastava, J.A. Kruijtzer, A.J. Heck, G.J. Boons, Chemoenzymatic approach for the preparation of asymmetric Bi-, tri-, and tetra-antennary N-Glycans from a common precursor, J. Am. Chem. Soc. 139 (2017) 1011—1018. https://doi.org/10.1021/jacs.6b12080.

[62] A. Hoshino, et al., Tumour exosome integrins determine organotropic metastasis, Nature 527 (2015) 329—335. https://doi.org/10.1038/nature15756.