

Ion mobility-tandem mass spectrometry of mucin-type *O*-glycans

Received: 27 October 2023

Accepted: 12 March 2024

Published online: 23 March 2024

Check for updates

Leila Bechtella ^{1,2}, Jin Chunsheng ³, Kerstin Fentker ^{1,4}, Güney R. Ertürk¹, Marc Safferthal ^{1,2}, Łukasz Polewski ^{1,2}, Michael Götze ^{1,2}, Simon Y. Graeber ^{5,6,7}, Gaël M. Vos^{1,2}, Weston B. Struwe ^{8,9}, Marcus A. Mall ^{5,6,7}, Philipp Mertins ^{4,10}, Niclas G. Karlsson^{3,11} & Kevin Pagel ^{1,2} ✉

The dense *O*-glycosylation of mucins plays an important role in the defensive properties of the mucus hydrogel. Aberrant glycosylation is often correlated with inflammation and pathology such as COPD, cancer, and Crohn's disease. The inherent complexity of glycans and the diversity in the *O*-core structure constitute fundamental challenges for the analysis of mucin-type *O*-glycans. Due to coexistence of multiple isomers, multidimensional workflows such as LC-MS are required. To separate the highly polar carbohydrates, porous graphitized carbon is often used as a stationary phase. However, LC-MS workflows are time-consuming and lack reproducibility. Here we present a rapid alternative for separating and identifying *O*-glycans released from mucins based on trapped ion mobility mass spectrometry. Compared to established LC-MS, the acquisition time is reduced from an hour to two minutes. To test the validity, the developed workflow was applied to sputum samples from cystic fibrosis patients to map *O*-glycosylation features associated with disease.

Mucins are the main component of mucus, a barrier-like hydrogel that is especially prominent in the enteric and pulmonary epithelia¹. They are large proteins that are heavily decorated with *O*-glycans, which constitute up to 80% of their molecular weight. Their glycosylation profile is highly heterogeneous and significantly altered in diseases such as cystic fibrosis (CF)², chronic obstructive pulmonary diseases (COPD), or cancer³, suggesting a role in disease progression and strong diagnostic potential. However, very little is known about the glycosylation molecular structures and the consequences on the glycan-dependent macromolecular interactions of the mucus. A fundamental challenge for the analysis of mucin-type *O*-glycans is the high

complexity of glycan structures; the constituting monosaccharides are often isomeric (e.g. GalNAc vs. GlcNAc), the building blocks can be linked in a variety of positions (e.g. 1-3 vs. 1-6), and the glycosidic bond can be oriented in different configurations (α vs. β anomers). This leads to highly branched and isomeric oligosaccharides, which often coexist but potentially have varying functional properties. These isomers cannot be easily distinguished by mass spectrometry alone⁴ and therefore orthogonal techniques are usually required for their identification. A second fundamental challenge arises from *O*-link core diversity⁵. In contrast to *N*-glycans, which share a conserved Man₃GlcNAc₂ core, eight different core structures are found for

¹Institute of Chemistry and Biochemistry, Freie Universität Berlin, Altensteinstraße 23A, 14195 Berlin, Germany. ²Fritz Haber Institute of the Max Planck Society, Faradayweg 4-6, 14195 Berlin, Germany. ³Department of Medical Biochemistry and Cell Biology, Institute of Biomedicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden. ⁴Max Delbrück Center for Molecular Medicine, Robert-Rössle-Str. 10, 13125 Berlin, Germany. ⁵Department of Pediatric Respiratory Medicine, Immunology and Critical Care Medicine and Cystic Fibrosis Center, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany. ⁶German Center for Lung Research (DZL), associated partner site, Berlin, Germany. ⁷Berlin Institute of Health at Charité – Universitätsmedizin Berlin, Berlin, Germany. ⁸Kavli Institute for Nanoscience Discovery, University of Oxford, Oxford OX1 3QU, UK. ⁹Department of Biochemistry, University of Oxford, Oxford OX1 3QU, UK. ¹⁰Berlin Institute of Health, 10178 Berlin, Germany. ¹¹Department of Life Sciences and Health, Faculty of Health Sciences, Oslo Metropolitan University, Oslo, Norway. ✉e-mail: kevin.pagel@fu-berlin.de

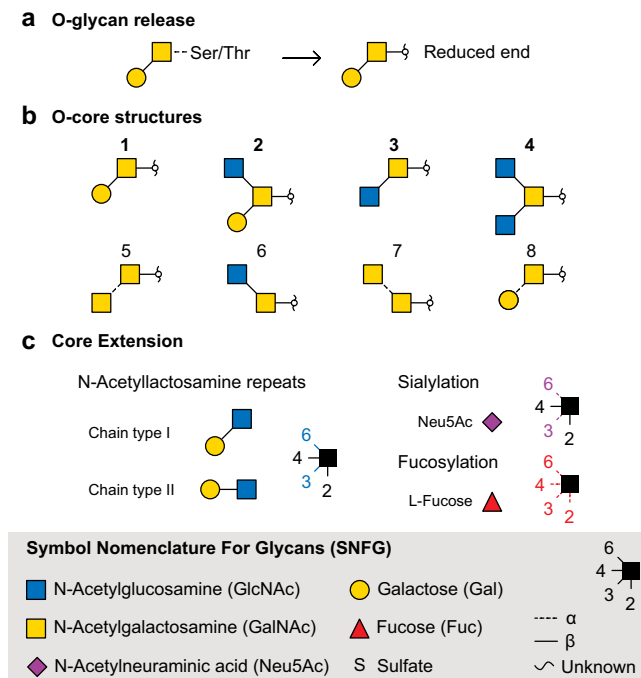


Fig. 1 | Typical O-glycan structures shown using the symbol nomenclature for glycans (SNFG)⁵⁹. **a** Representation of the O-glycosidic linkage to Ser/Thr and the alditol form resulting from β -elimination. **b** The most abundant O-glycan core structures 1–4 and the four less common cores 5–8. **c** Typical extensions of glycan cores are the attachment of N-acetylglucosamine repeats, sialylation and fucosylation.

O-glycans (Fig. 1). Mucin-type O-glycans are extended from a common GalNAc bound to the glycoprotein through the hydroxy group of a Ser or Thr side chain. As there is no universal enzyme able to cleave all O-glycan core structures collectively, chemical methods such as reductive β -elimination are usually employed. The released O-glycans are then converted to alditols, to prevent subsequent peeling reaction leading to the loss of the GalNAc at the reducing end^{6,7}, and dilution of the signal over the GalNAc α and β anomers⁸.

Typically, liquid chromatography (LC) coupled to tandem mass spectrometry (MS/MS) is employed for the analysis of mucin-type O-glycans^{8,9}. A stationary phase of porous graphitised carbon (PGC) is commonly used for the separation of non-derivatized O-glycans. The polar carbohydrates are well retained on the polarisable PGC phase¹⁰, while its planar surface adds a 3-dimensional separation¹¹. Although recognised for its ability to separate isomers, PGC tends to lack reproducibility^{12,13}. To maintain its separation properties, regular regeneration steps using acids or methanol are required throughout the analysis sequence^{12,13}. LC retention times are used as an additional level of identification for putative assignment of glycan structures, but these often vary significantly. Deviations in retention times are seemingly dependent on both the type of instrument as well as the employed experimental conditions, making inter-laboratory comparison difficult.

Following their separation, glycan isomers can be identified by tandem MS, typically via collision-induced dissociation (CID) of negatively charged ions. CID of deprotonated glycans leads to cross-ring fragments, which are diagnostic to the regiochemistry and core structure^{14–16}. Although highly informative, glycan MS/MS spectra are characteristically complex and difficult to interpret. An orthogonal structural technique to distinguish and identify isomeric structures with higher confidence is therefore highly desirable.

Ion mobility spectrometry (IMS) has recently emerged as a powerful technique to separate isomeric glycans in the gas phase^{17–19} In

traditional IMS devices, ions are propelled by an electric field through a mobility cell filled with an inert gas such as He or N₂. Low-energy collisions with the gas facilitate a temporal separation of ions according to their size, shape, and charge. The collision cross section (CCS) retrieved from mobility experiments is a molecular property and corresponds to the rotationally averaged area of the ion that collides with the drift gas²⁰. Depending on the type of instrument, CCSs can be measured directly or estimated after calibration. Unlike PGC-LC retention times, the CCS of an ion is instrument independent and can in principle be calculated theoretically. Furthermore, deprotonated glycans can be separated more efficiently in IMS than their protonated or sodiated forms¹⁸, while being less prone to rearrangement reactions such as fucose migration^{21–23}.

Recently, the potential of ion mobility spectrometry for the separation of O-glycan structures was tested and compared to PGC-LC separation²⁴. This work clearly demonstrated the potential for IMS to aid in the identification of otherwise challenging to assign O-glycan isomers through their CCS. However, the resolution achieved with the utilised travelling wave ion mobility spectrometer (TWIMS) was not sufficient to separate the majority of intact glycan ions and estimate their CCS. Instead, mobilities of diagnostic fragments had to be used to identify the parent structures.

The resolving power of traditional IMS techniques, such as drift tube IMS and TWIMS, is mostly governed by the dimensions of the drift cell. Recent advances in instrumentation address this issue either by performing cyclic IMS experiments in a closed-loop system²⁵, or by using distinct cell geometries and electric fields. In trapped ion mobility spectrometry (TIMS)²⁶, ions are driven by the drift gas through the cell and trapped by an electric field going in opposite direction. The progressive decrease of the electric field separates the ions according to their mobility. By manipulating the ramp of the electric field and the gas flow, the resolving power can be increased without altering the mobility cell.

Here, we present a TIMS-based method to separate and comprehensively characterise glycan isomers from complex biological samples. The complete analytical workflow takes only a few minutes, which makes it an attractive alternative to the rather slow LC-MS method and importantly improves IMS resolution as compared to aforementioned studies. Additionally, the measured CCSs enable a database-driven high-throughput O-glycan analysis.

Results

TIMS separation of O-glycan isomers

To validate the adequate resolving power of TIMS for natural mucin-type O-glycan analysis, the O-glycome of commercially available porcine gastric mucins (PGM) was investigated. PGM was selected for its highly heterogeneous but well-characterised O-glycosylation²⁴. The released glycan alditols were ionised through direct infusion and separated in the TIMS cell. Multiple IMS scans of 1 s ramp time were accumulated over two minutes to reach highest signal-to-noise ratio. In just two minutes, a detailed TIM-MS O-glycoprofile of PGM comparable to that of a 40 min PGC-LC-MS gradient was acquired (Fig. 2).

Most of the observed ions are singly deprotonated; doubly deprotonated species are only detected for glycans above 1200 Da. Figure 2a shows the 2D plot of the released PGM O-glycans with m/z on one the y-axis and $1/K_0$ on the x-axis. The highlighted region shows the trendlines of singly and doubly charged ions. The total mobilogram can be annotated in the same way as an LC chromatogram, using the CCS instead of the retention time in min (Fig. 2b). The general composition of the glycans was retrieved from the m/z values. Complete O-glycan structures were assigned from MS/MS spectra (for complete assignment see Supplementary Data 1). Forty-nine structures were identified in PGM, with in average two isomers per composition. Overall, the TIM-MS accumulated spectrum matches the PGC-LC-MS data well (Fig. 2c).

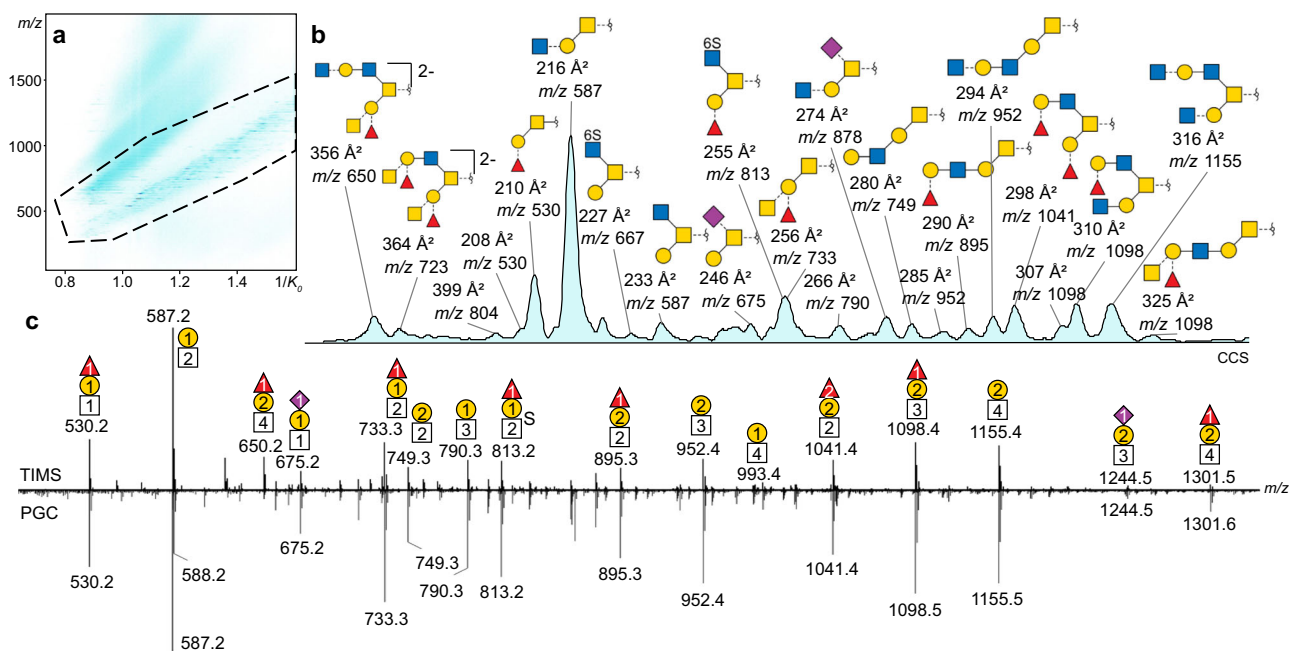


Fig. 2 | TIM-MS of deprotonated O-glycans from porcine gastric mucins. a TIM-MS m/z and $1/K_0$ heatmap with highlighted singly and doubly deprotonated O-glycan ions (dashed line). **b** Extracted ion mobilogram with collision cross section (CCS) and m/z of the most abundant species. **c** MS spectra accumulated over the

full ramp time/gradient in TIMS (2 min) and $P_{GC}LC$ (40 min). Glycan compositions and structures are shown using the SNFG nomenclature⁵⁹. The number of each residue is noted inside the symbol. All assigned structures, their CCSs, and their corresponding fragmentation spectra are given in the Supplementary Data 1.

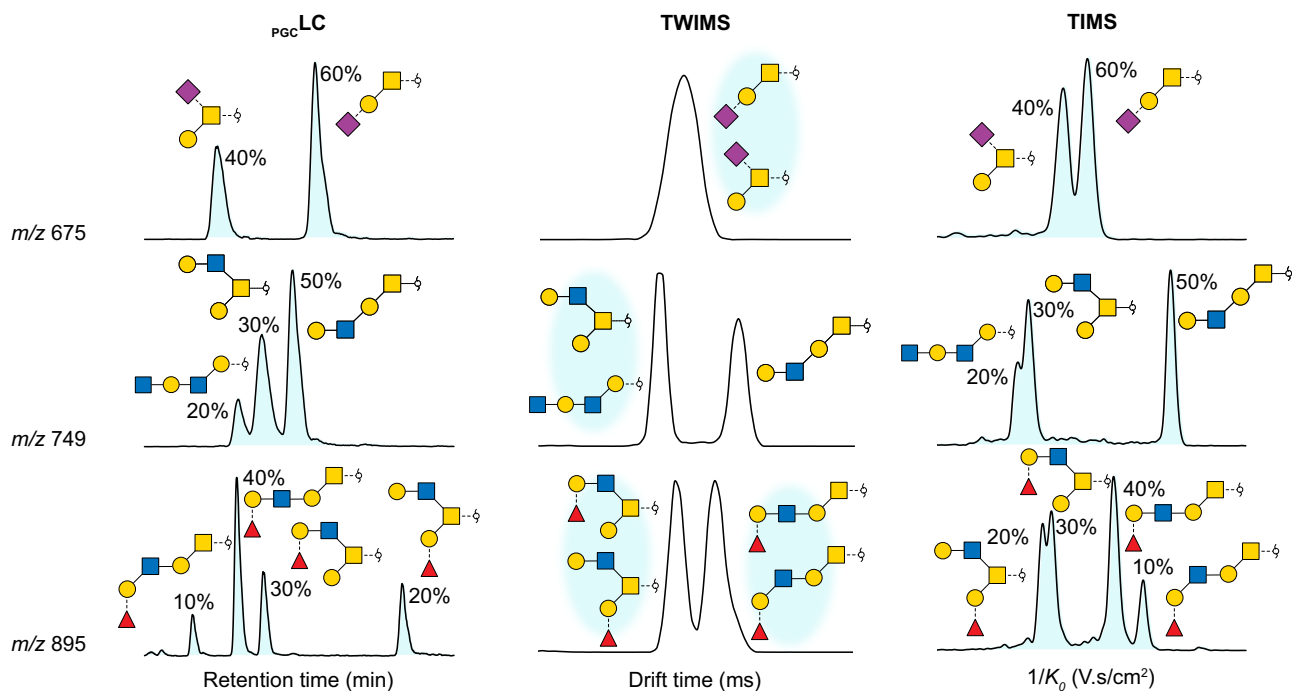


Fig. 3 | Efficiency of isomeric separation for a selection of three mucin-type O-glycans. The extracted ion chromatograms ($P_{GC}LC$) and mobilograms (TWIMS and TIMS) of deprotonated ions at m/z 675, 749 and 895 are shown. The relative area of each isomer is indicated in percentage.

Separation of glycan isomers is key for the structural characterisation of the O-glycome. $P_{GC}LC$ is a powerful method for O-glycan separation. However, it is also time-consuming and prone to error as retention time shifts can induce false putative assignments. TWIMS has previously also been tested for O-glycans analysis, albeit with limited resolving power²⁴. In order to assess the utility of the method presented here, the separation of selected deprotonated isomeric

structures using $P_{GC}LC$, TWIMS and TIMS are compared (Fig. 3). The relative peak area for each isomer is annotated in percentage of the sum area of isomeric structures at same m/z . While all O-glycan isomers are separated using $P_{GC}LC$, TWIMS clearly lacks the resolution to distinguish all isomers. TIMS on the other hand has a significantly improved resolution and sufficiently separates all structures that can be distinguished by $P_{GC}LC$. In addition, the relative peak areas of

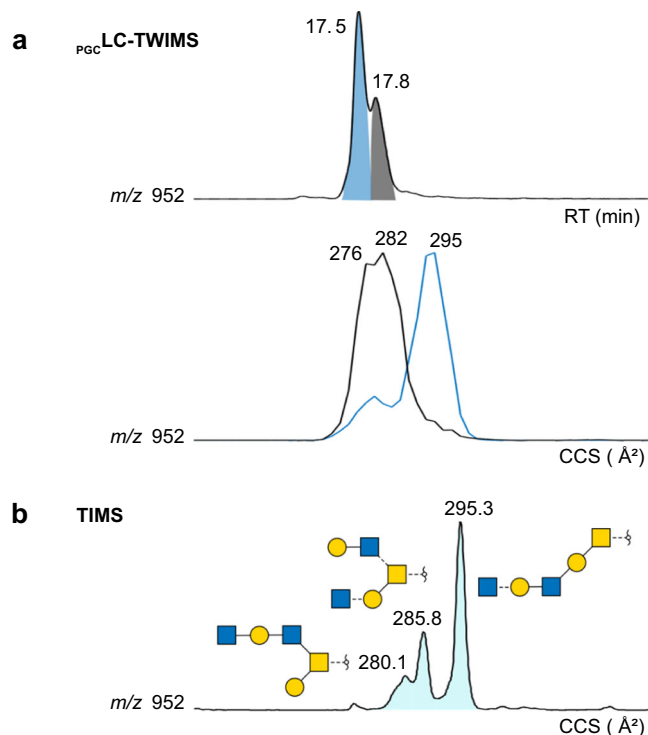


Fig. 4 | Separation of neutral isomeric *O*-glycans at m/z 952 using p_{GC} LC coupled to TWIMS or stand-alone TIMS. **a Extracted ion chromatogram of m/z 952 and corresponding ion mobilograms (TWIMS) for each p_{GC} LC peak. **b** Extracted ion mobilogram (TIMS) of m/z 952.**

selected structures do not differ drastically between p_{GC} LC and TIMS, which indicates that intensities can be compared for semi-quantification.

Some neutral structures are challenging to separate using p_{GC} LC and TWIMS (Fig. 4a). For both compositions, the direct hyphenation of p_{GC} LC and TWIMS reveals the existence of several isomers, but overlapping peaks prevent further characterisation. Due to the improved mobility resolution, TIMS can address this issue. All isomers can be resolved sufficiently (Fig. 4b), and the resulting fragmentation spectra allow the identification of their structures (SI).

TIM-MS/MS for glycans structural assignment

In a TIMS experiment, isomers are first separated in the mobility cell and subsequently analysed by MS/MS. The TIM cell is placed directly after the ionisation source, the mobility-separated ions are directed to the quadrupole for precursor ion selection, then to the collision cell for fragmentation and finally to the TOF for mass detection. As a result, the TIM-MS/MS workflow is conceptually comparable to that of LC-MS/MS, where an individual fragmentation spectrum can be obtained for each LC-separated species. This is shown in Fig. 5 where four isomers of m/z 895 are separated by TIMS and subsequently fragmented by MS/MS. The fragmentation of negatively charged glycans yields highly informative cross-ring fragments, from which the regiochemistry and the glycan core structure can be deduced^{15,16,27–29}. Here, diagnostic fragments indicate a core 2 for the structures of CCSs 273 Å² and 275 Å² and a core 1 for the structures of CCSs 290 Å² and 297 Å². An α -1,2 fucosylated Gal is found in all four isomers. A blood group H and a type II chain were attributed to the structures of CCS 275 Å² and 290 Å². The structure of CCS 297 Å² is a core 1 extended with a blood group H and a type I chain. *O*-glycan structures in PGM were systematically assigned and are all reported in the Supplementary Data 2. The results match well with previously reported structures²⁴.

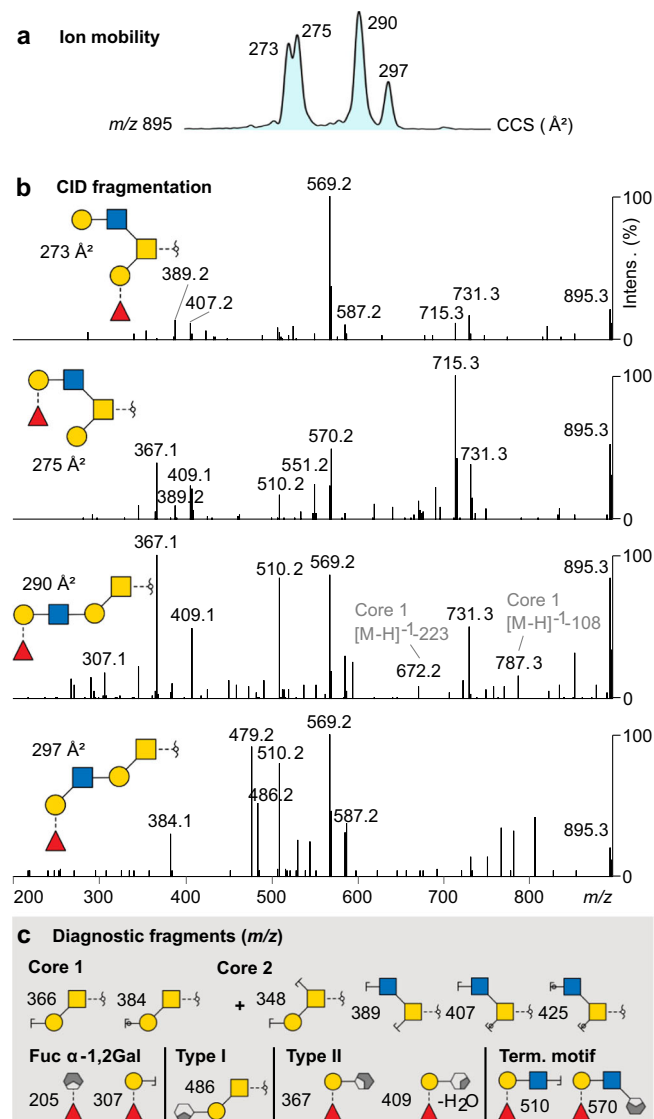


Fig. 5 | TIM-MS/MS structural assignment based on characteristic aldol fragments. **a Extracted TIM ion mobilogram of m/z 895. **b** Accumulated CID fragmentation spectra for each mobility peak of m/z 895 and corresponding identified structures. **c** Main characteristic fragments that aided the assignment of the four isomeric structures.**

CCS-based structural characterisation

Retention times are governed by multiple factors such as the physicochemical properties of the analyte and stationary phase. As a result, they often lack reproducibility. CCSs on the other hand are molecular properties that are directly linked to the structure of the ions in the gas phase. This makes CCS values particularly interesting for characterising isomers, where they can provide important information on the branching and regiochemistry of the investigated molecules.

Isomer separation by ion mobility can reveal minute structural differences. The ions at m/z 895 discussed above are a good example. The two structures of CCS 290 Å² and 297 Å² differ only by the presence of a β -1,4 or β -1,3 glycosidic bond, corresponding to type I and type II linkages, respectively (Fig. 5). This linkage detail is often challenging to identify by MS/MS alone and depends on the presence of diagnostic, but often low-abundant, cross-ring fragments. The type I structure yields a higher CCS than type II, which can be explained by the presence of the β -1,3 linkage inducing a more open structure. These observations can help assign similar structures more generally.

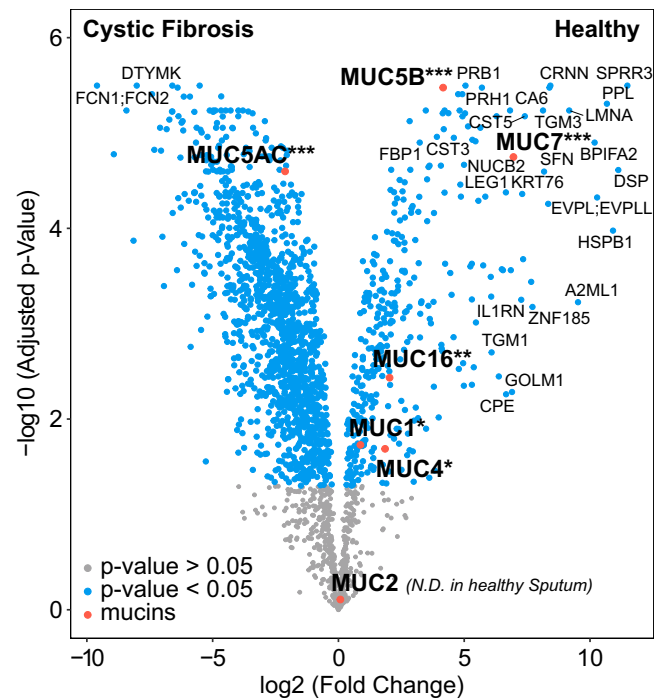
Aside from qualitative use, CCSs can be utilised for de novo annotations. CCSs are universally comparable when measured under controlled conditions³⁰. As a result, they can be readily implemented in online databases and used independently of the instrument and laboratory³¹. Direct CCS measurement can be performed on drift tube IMS instruments; other IMS approaches require calibration with calibrants of similar physico-chemical properties as the analyte to estimate CCS values. The CCSs of PGM *O*-glycans reported here were determined using two ion mobility spectrometers, TIMS and TWIMS, which use considerably different technologies to achieve mobility separation. The standard deviation between ^{TIMS}CCS_{N2} and ^{TWIMS}CCS_{N2} is consistently below 1% over the whole CCS range, which demonstrates the general validity of the approach. All CCS values and the corresponding STDs between the two instruments are given in the Supplementary Data 2. The processed data is available on the online database UniCarb-DR, with CCS values indicated in the place of retention times. The accession codes are provided in the Data Availability section.

Clinical sputum *O*-glycan profiling

As a proof-of-principle, the TIM-MS/MS method was applied to two clinical samples of healthy and cystic fibrosis (CF) sputum. Patients with CF produce a substantial amount of sputum as the disease is associated with a hypersecretion of mucus. For this reason, the glycosylation of secreted mucins in CF sputum has been thoroughly investigated. The stage of the muco-obstructive lung disease induces drastic variations in the mucus *O*-glycan profile, as previously reported^{2,32–34}. Here, healthy and CF sputum samples are used to verify if the TIM-MS/MS method can discriminate the two clinical samples and highlight prominent glycosylation alterations in CF sputum. First, the relevance of the samples to the disease state was assessed using proteomics. Of the 2283 quantified proteins, 1375 have a significantly different abundance (adj. *p*-value < 0.05). The complete protein identification output is reported in Supplementary Data 3. The main contributors to glycosylation in the sputum with a molecular weight above 500 kDa are mucins. Therefore, we compared the log₂ intensities of the four quantified high molecular weight mucins (Fig. 6). MUC5B is the most abundant mucin in both samples and therefore likely contributes the most to the glycosylation profiles. In the healthy sample, the contribution is higher than in the CF sample. MUC5AC is the second most abundant mucin and higher abundant in the CF sputum. MUC2 has only been quantified in the CF sample and MUC1 contributes the least to the glycosylation because of its low abundance. Overall, the data shows that the sputum samples contain MUC4, MUC1, MUC16, MUC7, MUC5AC and MUC5B. The increased level of MUC5AC and decreased level of MUC5B in CF has been reported previously^{35,36}. The proteomics data therefore confirm that the chosen samples follow the general trends observed in CF samples. As size exclusion chromatography with a mass cut-off of 600 kDa was used to collect the mucin fraction before performing glycomics, we assume that the sputum samples used for glycomics contain mostly MUC5AC and MUC5B.

TIM-MS/MS analysis reveals strong differences in the *O*-glycosylation profile of healthy and CF sputum (Fig. 7a, for complete assignment, see Supplementary Data 1 and 2). In general, the glycans are much simpler in the CF sample and decorated to a higher degree with sialic acid. The diversity in glycan structures is reduced in the CF sample, for which the mobilogram area seem to be concentrated on a few peaks of high intensity. This agrees well with previous studies revealing a generally reduced complexity in CF glycosylation².

To compare the relative abundance of core structures and specific glycan features such as sialylation, fucosylation and sulfation, the mobility peak areas have been integrated and combined into groups (Fig. 7b, c). The CF sample presents less core 3 and core 4 compared to the healthy one, reflecting a significant loss of diversity



Mucin	MW (kDa)	Log ₂ (LFQ)		
		CF	Healthy	Adj. p-Value
MUC5AC	~2200; >500	29.8	27.7	0.000 ***
MUC2	~550 or higher	24.1	<i>N.D.</i>	
MUC1	200-500	25.7	26.6	0.194 <i>ns</i>
MUC4	320	24.0	25.9	0.252 <i>ns</i>
MUC16	>250	25.1	27.2	0.033 *
MUC7	150-200	24.2	31.1	0.043 *
MUC5B	1000-2000	30.4	34.6	0.000 ***

Fig. 6 | Comparison of protein abundances between healthy and cystic fibrosis sputum. Proteins of significantly different intensities (adjusted *p*-value < 0.05, moderated *t*-test) are highlighted in blue. Additionally, mucin proteins are highlighted in red. The table indicates mean Log₂ label-free quantification (LFQ) intensities and significance for each mucin calculated using a two-sided *t*-test. A Benjamini–Hochberg method was used to adjust for multiple comparisons. *** adj. *p* < 0.001, ** adj. *p* < 0.01, * adj. *p* < 0.05. The molecular weight is indicated for each mucin^{60–70}. Source data are provided as a Source Data file.

in the CF core structures. Sialylated glycans are three times more abundant in the CF sputum, while fucosylation and sulfation are decreased. Similar trends were reported in the past using *PGC*-LC-MS on purified MUC5B and MUC5AC^{32,37}. Specifically, the high level of sialylation in CF sputum and inflammatory state is consensus of various studies^{32,34,37,38}.

Closer inspection of the relative abundance of isomeric structures provides a deeper level of information (Fig. 8). Even when all isomeric structures are present in the sample, the peak ratios can identify differences in their abundance. In case of the trisaccharides at *m/z* 587, these changes are notable, with a complete inversion in abundance. A simple qualitative analysis of the samples would not reveal such differences.

The TIM-MS method reported here can provide *O*-glycan profiles from clinical samples in as little as two minutes. In addition, the CCSs of all identified *O*-glycans were determined and can be used as references

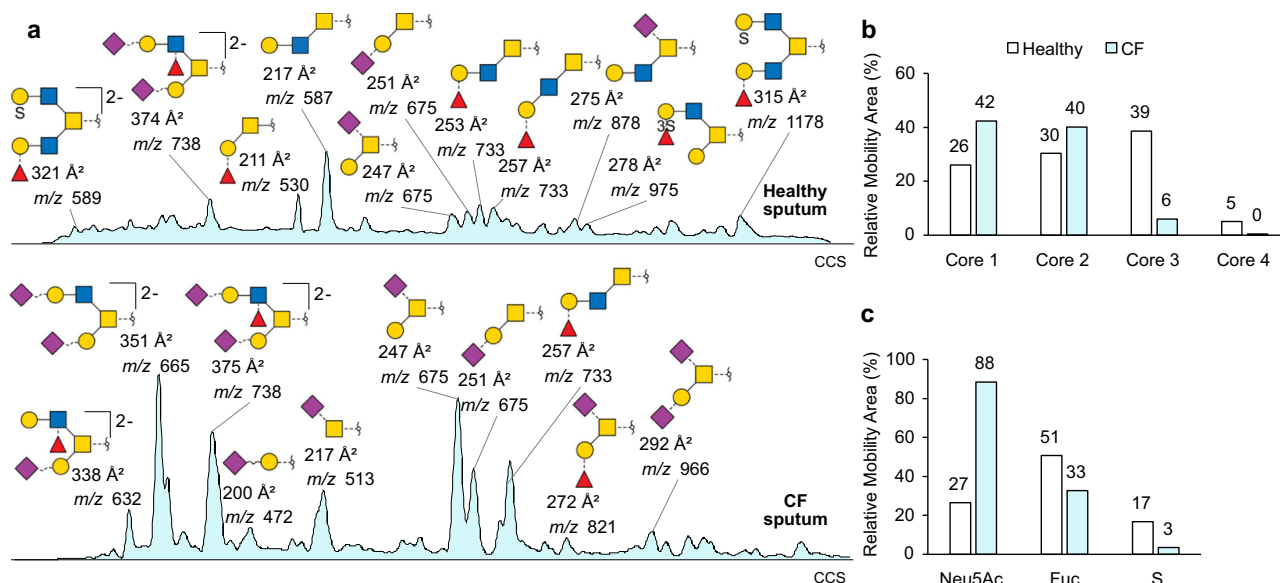


Fig. 7 | O-glycosylation profile of clinical sputum by TIM-MS/MS. a Glycosylation profiles of healthy sputum (top) and cystic fibrosis (CF) sputum (bottom) obtained by TIM-MS/MS. The content in **b** core structures and **c** glycosylation features (sialylation, Neu5Ac; fucosylation, Fuc; and sulfation, S) are shown in relative

mobility areas. The detailed fragmentation data is reported in Supplementary Data 1 and all assigned structures with their corresponding CCS values and mobility peak area are reported in the Supplementary Data 2.

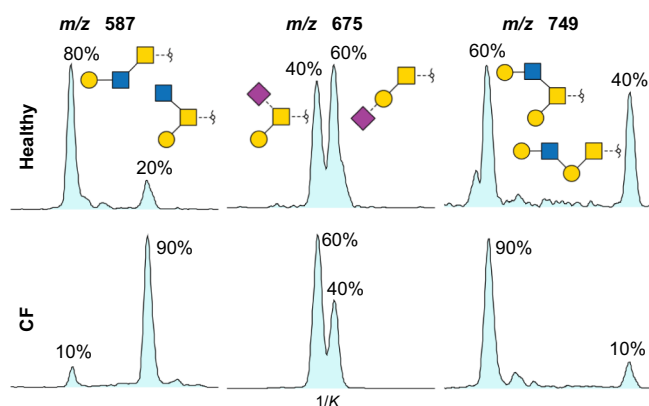


Fig. 8 | Abundance of isomeric structures in healthy and cystic fibrosis (CF) samples. Extracted ion mobilograms of m/z 587, 675 and 749 (from left to right) in healthy (top) and CF (bottom) sputum. The relative area of each isomer is indicated in percentage.

for future assignment. This combination of short analysis time and diagnostic CCSs enables a high-throughput analysis of *O*-glycosylation features in clinical samples.

Discussion

For the analysis of complex samples, it is crucial to separate their constituents, especially when multiple isomers coexist as in the case of glycans. Liquid chromatography can profile *O*-glycans in approximately one hour. The work presented here shows that ion mobility spectrometry is much faster and provides a qualitatively similar overview in just a couple of minutes, which makes it a powerful alternative to classical LC-MS workflows.

Until now, the use of ion mobility in the omics fields is largely based on implementation into existing LC-MS workflows as a third dimension. Here, we provide a method that utilises IMS as an independent, stand-alone separation technique for the analysis of complex, highly heterogeneous samples. We show that the stand-alone

TIMS technology with the well-tuned mobility parameters leads to mobilograms that in terms of resolution and informational content resemble LC chromatograms. In addition to the rapid separation, IMS provides direct structural information in the form of collision cross sections, which can be utilised for structural assignment in complex omics analyses.

The structural complexity of glycans is still the major bottleneck for their analysis. Recent advancements in ion mobility technology highlighted the potential of the method to be used in LC-based glycomics analyses^{24,39–44}. Fundamental studies recently showed that IMS can be theoretically described using a plate-height model, similarly to other separation techniques such as capillary zone electrophoresis or LC^{45,46}. Dedicated bioinformatics tools for deconvolution and data treatment further strengthen the utility of ion mobility in real-world applications^{31,47,48}. Finally, also the often tedious fragmentation-based assignment of *O*-glycans structures is currently on the verge of automation, for example, via deep-learning approaches⁴⁹. When combined, these advancements have the potential to transform glycomics from a highly specialized technology into a broader, high-throughput application.

Methods

Ethical statement

This research complies with all relevant ethical regulations. The collection of sputum samples was approved by the ethics committee of Charité - Universitätsmedizin Berlin (EA2/O16/18). Written informed consent was obtained from all participants.

Sputum collection

Two sputum samples were used in this study: one of a healthy donor and one of a cystic fibrosis patient, and consent has been obtained for sharing of individual-level data. Samples were matched for sex and gender, both female, and age at the time of collection, was between 28 and 33. Sex and gender were therefore not considered in the study design. The healthy sample was collected after induction with inhaled hypertonic saline (NaCl 6%) and the CF sample was collected after spontaneous expectoration. Samples were stored at -80°C until analysis.

O-glycan sample preparation

Porcine gastric mucin (PGM) was obtained from Sigma-Aldrich (U.S.A.). The clinical mucus samples were first heat-inactivated. In the same step, samples were reduced and alkylated, and detergent was added to the mixture to disrupt the mucus hydrogel. Briefly, an equal volume of buffer containing 4% SDS, 100 mM Tris-HCl pH 8, 1 mM EDTA, 150 mM NaCl, 20 mM DTT, 80 mM CAA was added to the mucus. The mixture was heated at 95 °C for 10 min. After cooling down, the samples were incubated with 25 U of benzonase for 15 min at room temperature. The mucins fraction of the samples was purified by size exclusion chromatography (SEC) using an AZURA Multi Method FPLC System (Knauer, Germany). Superdex 200 (prep grade, Cytiva, USA) was used to pack a column of 37 mL bed volume. The separation was performed on 200 μ L of inactivated sample, at a flow rate of 0.75 mL/min in 100 mM ammonium acetate. Absorbance was recorded at 214 nm. The mucin fraction was collected in the void volume, lyophilised, and resuspended in 200 μ L H₂O. O-glycans from PGM (2 mg/mL in H₂O) and from the collected sputum mucins were released by reductive β -elimination⁸. Samples (160 μ L) were incubated with 20 μ L 5 M NaBH₄ and 20 μ L 0.5 mM NaOH at 50 °C for 16 h. The reaction was quenched with 20 μ L acetic acid and desalted on 400 mg Dowex 50WX8 cation exchange beads (Sigma-Aldrich, USA). The resin was first washed three times with methanol (MeOH), and conditioned with 1 mL HCl 1 M, 1 mL MeOH and 1 mL H₂O. Glycans were loaded on the resin and eluted with two times 500 μ L H₂O. O-glycans were then enriched on a 50 mg Hypercarb SPE cartridge (Thermo Fischer). The cartridge was conditioned with 1 mL ACN 0.1% TFA, and 1 mL H₂O 0.1% TFA. Glycans were loaded on the cartridge and washed three times with 1 mL H₂O 0.1% TFA. Glycans were eluted with four times 100 μ L of 50% ACN 0.1% TFA and dried in a SpeedVac.

O-glycan analysis by PGC-LC-MSⁿ and TWIMS

PGC-LC-MS, PGC-LC-MS/MS and PGC-LC-TWIMS were performed using a SYNAPT G2-Si spectrometer (Waters, Manchester, U.K.) equipped with an Acquity UPLC system. O-glycan alditols were separated at room temperature using a 100 \times 2.1 mm I.D. PGC column of 5 μ m particle size (Hypercarb, Thermo Scientific, U.S.A.). The released O-glycans were dissolved in 40 μ L H₂O and 5 μ L were injected. Glycans were eluted using a linear gradient from 0 to 40% acetonitrile (ACN) in 10 mM NH₄HCO₃ over 40 min at a flow rate of 150 μ L/min. Glycan alditols were ionised in ESI negative ion mode with a capillary voltage of 2.8 kV, and the source temperature was 150 °C. The mass range was set to m/z 50–3500. Fragmentation was performed using ramps of collision energy dependent on the m/z ratio of the precursor ions, in the range of 15 to 120 eV. O-glycan mobilities were recorded after LC separation in a separate acquisition, using an IMS wave velocity of 450 m/s, and an IMS wave height of 40 V. The data was acquired and processed using the MassLynx software (version 2.0.7, Waters). To obtain relative peak areas, the chromatograms were deconvoluted and integrated with MZmine 2⁵⁰. O-glycans structures were manually assigned from the MS/MS spectra, based on their diagnostic fragment ions, using GlycoWorkBench 2.1⁵¹.

O-glycan analysis by TIM-MSⁿ

TIM-MS and TIM-MS/MS were performed on a timsTOF Pro spectrometer (Bruker, Bremen, Germany), equipped with an in-house 3D-printed offline nano-ESI source, whose design was recently published⁵². Released glycans were dissolved in 50 mM ammonium acetate in H₂O:MeOH (1:1). For each measurement, 5 μ L of sample were introduced in a Pt/Pd-coated glass capillary emitter prepared in-house and ionised in negative ion mode nano-ESI. The instrument parameters were set as follows: capillary voltage 1.2 kV, end plate offset -0.5 kV, dry source temperature 150 °C, D1 = 20 V, D2 = 30 V, D3 = -100 V, D4 = -450 V, D5 = 60 V, and D6 = -46 V. The TIMS separation was performed in

nitrogen (N₂). Ions reversed mobility $1/K_0$ was scanned between 0.6 V·s/cm² and 1.6 V·s/cm². The accumulation time and the ramp time were set to 20 ms and 1000 ms, respectively. The transfer time was fixed to 2 ms. Ions were selected in the quadrupole after mobility separation and fragmented in the collision cell, with collision energy varying from 25 to 120 eV. The data was acquired with Compass otofControl (version 6.2, Bruker). The mobilograms were processed and integrated using DataAnalysis (version 4.0, Bruker). O-glycans structures were manually assigned using GlycoWorkBench 2.1⁵¹.

Mobility calibration and data treatment

TWIMS arrival time distributions were first fitted to a Gaussian distribution with Origin 95E before estimating ^{TWIMS}CCSs. TWIMS calibration was performed based on a described protocol⁵³, using a dextran ladder 1K (Sigma-Aldrich) of known drift tube CCS in N₂, ^{DT}CCS_{N₂}^{54,55}. Briefly, a logarithmic calibration curve was plotted using the corrected drift time t_D' (in ms) and corrected CCS' (in Å²), using the following equations:

$$t_D' = t_D - C \frac{\sqrt{m/z}}{1000}, \quad (1)$$

where t_D is the experimental drift time and C is the EDC Delay Coefficient (1.41).

$$\text{CCS}' = \text{CCS} \times \frac{\sqrt{\mu}}{z}, \quad (2)$$

where μ is the reduced mass.

The fitted drift time t_D was corrected by subtracting the travel time from the end of the mobility cell to the entrance of the mass spectrometer in (1). The reference CCS value was corrected by the reduced mass and charge in (2).

As TIMS experiments result in the temporal release of the ions, dependent on the electric field gradient, and not in a drift time, a calibration is needed to convert the ion transit time t_t in the ion mobility cell into a reversed mobility $1/K_0$ ²⁶. This was achieved through the otofControl software (Bruker), using the ESI-Low-concentration tuning mix from Agilent. TIMS $1/K_0$ values of the dextran ladder 1K were then recorded and plotted against their corrected CCS' (in Å²), using Eq. (2). This external correction of the TIMS mobilities led to better results in terms of inter-instruments reproducibility of the CCS_{N₂}.

All ^{TWIMS}CCS_{N₂} and ^{TIMS}CCS_{N₂} values are listed in the Supplementary Data 2, which presents averages of two or three replicates acquired in independent measurements on each instrument.

Proteomics of clinical samples

Sputum samples were inactivated in SDS buffer (2% SDS, 50 mM Tris-HCl pH 8, 0.5 mM EDTA) at 95 °C for 10 min. After measuring the protein concentration using BCA assay, 100 μ g protein were reduced and alkylated in SDS buffer with 20 mM DTT, 40 mM CAA and 75 mM NaCl at 95 °C for 10 min. Subsequently, the mixture was incubated with 25 U benzonase for 15 min at room temperature and proteins were extracted using an SP3 protocol described for proteomics⁵⁶. Sera-Mag SpeedBeads (GE Healthcare, cat. no. 45152105050250 and 65152105050250) were prepared in a 1:1 mix at 50 mg/mL. Proteins were precipitated in 70% ACN and incubated with 5 μ L of the beads mix, at 1 μ g protein:2.5 μ g beads, for 20 minutes under shaking (800 rpm). The beads were washed twice with 70% ethanol and once with 100% ACN on a magnetic rack. After removal of the ACN, beads were resuspended in 50 mM ABC buffer and incubated with PNGase F for 1 h at 37 °C. Trypsin and LysC were added at 1:50 (enzyme:substrate) for overnight digestion. The supernatant was collected and

desalted on C18 material. The analysis was performed using an EASY-nLC 1200 System coupled to an Orbitrap Exploris 480 mass spectrometer (Thermo Fisher Scientific). Peptides (2 μ L at 1 μ g/ μ L) were injected two and three times for healthy and CF samples, respectively. Peptides were separated at 45 °C using a 20 cm reverse-phase column (Fused Silica Capillary I.D. 75 μ m, CM Scientific), packed in-house with 1.9 μ m C18-reprosil beads (r119.aq.0003, Dr. Maisch). Eluents A (3% ACN, 0.1% FA) and B (90% ACN, 0.1% FA) were used. Peptides were eluted at a flow rate of 250 nL/min, applying the following 200 min gradient: 2% to 3% B over 5 min, then from 3% to 20% B over 157 min, to 30% B over 30 min, to 60% B over 10 min, and finally to 90% over 1 min. The data was acquired using a top20 method in positive data-dependent acquisition mode using a Nanospray Flex™ ion source ES072, at spray voltage 2000 V, and ion transfer tube temperature 275 °C. The mass range was set to m/z 350–1600. The MS/MS spectra were generated using a normalised collision energy at 28. The unassigned charge states and charge states 1, 7, or higher were excluded from fragmentation. Dynamic exclusion was set to 20 s. The data was acquired using Thermo EASY-nLC 1200: 3.2.0 SP2 (version 4.1.4.1), Orbitrap Exploris 480 (tune software version 3.1.279.9) and Xcalibur (version 4.6.67.17). For protein identification, MaxQuant software (Version 2.0.3.0; Max Planck Institute of Biochemistry) and a decoy human UniProt database (2022-01) were used. The variable modifications oxidation (M), N-terminal acetylation, deamidation (N, Q) and the fixed modification carbamidomethyl cysteine were considered in the search. Peptides and proteins, which were identified with a false discovery rate of 1% were used for quantification. For quantification, match between runs and label-free quantitation (LFQ) algorithms were used. Downstream analysis was done in R (Version 4.0.4). Protein groups were filtered by reverse hits, proteins only identified by site, potential contaminants and for valid values in at least 60% of the samples. LFQ values were log₂ transformed and missing values were replaced by random values from a normal distribution with a width of 0.3 and a downshift of 1.8. For statistical testing moderated t-test (limma package⁵⁷) and two-sided t-test (rstatix⁵⁸) with Benjamini–Hochberg correction for multiple comparisons were used.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The glycomics and proteomics raw data generated in this study have been deposited to the ProteomeXchange Consortium via the PRIDE repository under the dataset identifier [PXD050530](#). The processed glycomics data generated in this study, including O-glycan structures, CCSs and MS/MS spectra, have been deposited in the UniCarb-DR database under accession codes: [533](#) (PGM); [534](#) (healthy sputum); [535](#) (CF sputum). The processed glycomics data generated in this study, including the annotated MS/MS spectra of O-glycan structures identified, are provided in the Supplementary Data 1. The processed glycomics data generated in this study, including the identified O-glycan structures with CCS, putative structure, area ratios from _{PGC}LC-TWIMS and TIMS data are provided in the Supplementary Data 2. The processed proteomics data generated in this study, including protein intensities reported by the MaxQuant identification software, are provided in the Supplementary Data 3. Source data are provided with this paper. The source data underlying Fig. 6 are provided as a Source Data file. Source data are provided with this paper.

References

1. Bansil, R. & Turner, B. S. Mucin structure, aggregation, physiological functions and biomedical applications. *Curr. Opin. Colloid Interface Sci.* **11**, 164–170 (2006).
2. Venkatakrishnan, V., Packer, N. H. & Thaysen-Andersen, M. Host mucin glycosylation plays a role in bacterial adhesion in lungs of individuals with cystic fibrosis. *Expert Rev. Respir. Med.* **7**, 553–576 (2013).
3. Kailemia, M. J., Park, D. & Lebrilla, C. B. Glycans and glycoproteins as specific biomarkers for cancer. *Anal. Bioanal. Chem.* **409**, 395–410 (2017).
4. Grabarics, M. et al. Mass spectrometry-based techniques to elucidate the sugar code. *Chem. Rev.* **122**, 7840–7908 (2022).
5. Karlsson, N. G., Jin, C., Rojas-Macias, M. A. & Adamczyk, B. Next generation O-linked glycomics. *Trends Glycosci. Glycotechnol.* **29**, E35–E46 (2017).
6. Mulagapati, S., Koppolu, V. & Raju, T. S. Decoding of O-linked glycosylation by mass spectrometry. *Biochemistry* **56**, 1218–1226 (2017).
7. Wilkinson, H. & Saldova, R. Current methods for the characterization of O-glycans. *J. Proteome Res.* **19**, 3890–3905 (2020).
8. Jensen, P. H., Karlsson, N. G., Kolarich, D. & Packer, N. H. Structural analysis of N- and O-glycans released from glycoproteins. *Nat. Protoc.* **7**, 1299–1310 (2012).
9. Schulz, B. L., Packer, N. H. & Karlsson, N. G. Small-scale analysis of O-linked oligosaccharides from glycoproteins and mucins separated by gel electrophoresis. *Anal. Chem.* **74**, 6088–6097 (2002).
10. Kolarich, D., Windwarder, M., Alagesan, K. & Altmann, F. *Glyco-Engineering: Methods and Protocols* (ed. Castilho, A.) vol. 1321 427–435 (Springer New York, 2015).
11. Melmer, M., Stangler, T., Premstaller, A. & Lindner, W. Comparison of hydrophilic-interaction, reversed-phase and porous graphitic carbon chromatography for glycan analysis. *J. Chromatogr. A* **1218**, 118–123 (2011).
12. Bapiro, T. E., Richards, F. M. & Jodrell, D. I. Understanding the Complexity of Porous Graphitic Carbon (PGC) chromatography: modulation of mobile-stationary phase interactions overcomes loss of retention and reduces variability. *Anal. Chem.* **88**, 6190–6194 (2016).
13. Pereira, L. Porous graphitic carbon as a stationary phase in HPLC: theory and applications. *J. Liq. Chromatogr. Relat. Technol.* **31**, 1687–1731 (2008).
14. Harvey, D. J. Fragmentation of negative ions from carbohydrates: Part 1. Use of nitrate and other anionic adducts for the production of negative ion electrospray spectra from N-linked carbohydrates. *J. Am. Soc. Mass Spectrom.* **16**, 622–630 (2005).
15. Everest-Dass, A. V., Abrahams, J. L., Kolarich, D., Packer, N. H. & Campbell, M. P. Structural feature ions for distinguishing N- and O-linked glycan isomers by LC-ESI-IT MS/MS. *J. Am. Soc. Mass Spectrom.* **24**, 895–906 (2013).
16. Karlsson, N. G., Schulz, B. L. & Packer, N. H. Structural determination of neutral O-linked oligosaccharide alditols by negative ion LC-electrospray-MSn. *J. Am. Soc. Mass Spectrom.* **15**, 659–672 (2004).
17. Hofmann, J. & Pagel, K. Glycan analysis by ion mobility–mass spectrometry. *Angew. Chem. - Int. Ed.* **56**, 8342–8349 (2017).
18. Hofmann, J., Hahm, H. S., Seeberger, P. H. & Pagel, K. Identification of carbohydrate anomers using ion mobility–mass spectrometry. *Nature* **526**, 241–244 (2015).
19. May, J. C. et al. Conformational ordering of biomolecules in the gas phase: nitrogen collision cross sections measured on a prototype high resolution drift tube ion mobility–mass spectrometer. *Anal. Chem.* **86**, 2107–2116 (2014).
20. Gabelica, V. et al. Recommendations for reporting ion mobility Mass Spectrometry measurements. *Mass Spectrom. Rev.* **38**, 291–320 (2019).
21. Lettow, M. et al. The role of the mobile proton in fucose migration. *Anal. Bioanal. Chem.* **411**, 4637–4645 (2019).

22. Mucha, E. et al. Fucose migration in intact protonated glycan ions: a universal phenomenon in mass spectrometry. *Angew. Chem.* **130**, 7562–7565 (2018).
23. Lettow, M. et al. Decoding the fucose migration product during mass-spectrometric analysis of blood group epitopes. *Angew. Chem. Int. Ed.* **62**, e202302883 (2023).
24. Jin, C., Harvey, D. J., Struwe, W. B. & Karlsson, N. G. Separation of isomeric O-glycans by ion mobility and liquid chromatography–mass spectrometry. *Anal. Chem.* **91**, 10604–10613 (2019).
25. Giles, K. et al. A cyclic ion mobility-mass spectrometry system. *Anal. Chem.* **91**, 8564–8573 (2019).
26. Michelmann, K., Silveira, J. A., Ridgeway, M. E. & Park, M. A. Fundamentals of trapped ion mobility spectrometry. *J. Am. Soc. Mass Spectrom.* **26**, 14–24 (2014).
27. Harvey, D. J. Fragmentation of negative ions from carbohydrates: Part 2. Fragmentation of high-mannose N-linked glycans. *J. Am. Soc. Mass Spectrom.* **16**, 631–646 (2005).
28. Harvey, D. J. Fragmentation of negative ions from carbohydrates: Part 3. Fragmentation of hybrid and complex N-linked glycans. *J. Am. Soc. Mass Spectrom.* **16**, 647–659 (2005).
29. Harvey, D. J., Royle, L., Radcliffe, C. M., Rudd, P. M. & Dwek, R. A. Structural and quantitative analysis of N-linked glycans by matrix-assisted laser desorption/ionization and negative ion nanospray mass spectrometry. *Anal. Biochem.* **376**, 44–60 (2008).
30. Stow, S. M. et al. An interlaboratory evaluation of drift tube ion mobility–mass spectrometry collision cross section measurements. *Anal. Chem.* **89**, 9048–9055 (2017).
31. Struwe, W. B., Pagel, K., Benesch, J. L. P., Harvey, D. J. & Campbell, M. P. GlycoMob: an ion mobility-mass spectrometry collision cross section database for glycomics. *Glycoconj. J.* **33**, 399–404 (2016).
32. Schulz, B. L. et al. Glycosylation of sputum mucins is altered in cystic fibrosis patients. *Glycobiology* **17**, 698–712 (2007).
33. Caballero, I. et al. Evidence of early increased sialylation of airway mucins and defective mucociliary clearance in CFTR-deficient piglets. *J. Cyst. Fibros.* **20**, 173–182 (2021).
34. Xia, B., Royall, J. A., Damera, G., Sachdev, G. P. & Cummings, R. D. Altered O-glycosylation and sulfation of airway mucins associated with cystic fibrosis. *Glycobiology* **15**, 747–775 (2005).
35. Henderson, A. G. et al. Cystic fibrosis airway secretions exhibit mucin hyperconcentration and increased osmotic pressure. *J. Clin. Investig.* **124**, 3047–3060 (2014).
36. Schaupp, L. et al. Longitudinal effects of elxacaftor/tezacaftor/ivacaftor on sputum viscoelastic properties, airway infection and inflammation in patients with cystic fibrosis. *Eur. Respir. J.* **62**, 2202153 (2023).
37. Hayes, C. A., Nemes, S. & Karlsson, N. G. Statistical analysis of glycosylation profiles to compare tissue type and inflammatory disease state. *Bioinformatics* **28**, 1669–1676 (2012).
38. Davril, M. et al. The sialylation of bronchial mucins secreted by patients suffering from cystic fibrosis or from chronic bronchitis is related to the severity of airway infection. *Glycobiology* **9**, 311–321 (1999).
39. Manz, C. et al. Determination of sialic acid isomers from released N-glycans using ion mobility spectrometry. *Anal. Chem.* **94**, 13323–13331 (2022).
40. Sastre Toraño, J. et al. Ion-mobility spectrometry can assign exact fucosyl positions in glycans and prevent misinterpretation of mass-spectrometry data after gas-phase rearrangement. *Angew. Chem.* **131**, 17780–17784 (2019).
41. Peterson, T. L. & Nagy, G. Toward sequencing the human milk glycome: high-resolution cyclic ion mobility separations of core human milk oligosaccharide building blocks. *Anal. Chem.* **93**, 9397–9407 (2021).
42. Pallister, E. G. et al. Utility of ion-mobility spectrometry for deducing branching of multiply charged glycans and glycopeptides in a high-throughput positive ion LC-FLR-IMS-MS workflow. *Anal. Chem.* **92**, 15323–15335 (2020).
43. Manz, C., Götze, M., Frank, C., Zappe, A. & Pagel, K. Dextran as internal calibrant for N-glycan analysis by liquid chromatography coupled to ion mobility-mass spectrometry. *Anal. Bioanal. Chem.* **414**, 5023–5031 (2022).
44. Vos, G. M. et al. Sialic acid O-acetylation patterns and glycosidic linkage type determination by ion mobility-mass spectrometry. *Nat. Commun.* **14**, 6795 (2023).
45. Grabarics, M. et al. Plate-height model of ion mobility-mass spectrometry. *Analyst* **145**, 6313–6333 (2020).
46. Grabarics, M. et al. Plate-height model of ion mobility-mass spectrometry: Part 2—Peak-to-peak resolution and peak capacity. *J. Sep. Sci.* **44**, 2798–2813 (2021).
47. Schmid, R. et al. Integrative analysis of multimodal mass spectrometry data in MZmine 3. *Nat. Biotechnol.* **41**, 447–449 (2023).
48. Ollivier, S., Fanuel, M., Rogniaux, H. & Ropartz, D. Molecular networking of high-resolution tandem ion mobility spectra: a structurally relevant way of organizing data in glycomics? *Anal. Chem.* **93**, 10871–10878 (2021).
49. Urban, J. et al. Predicting glycan structure from tandem mass spectrometry via deep learning. *bioRxiv* <https://doi.org/10.1101/2023.06.13.544793> (2023).
50. Pluskal, T., Castillo, S., Villar-Briones, A. & Orešič, M. MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinform.* **11**, 395 (2010).
51. Ceroni, A. et al. GlycoWorkbench: a tool for the computer-assisted annotation of mass spectra of glycans. *J. Proteome Res.* **7**, 1650–1659 (2008).
52. Götze, M., Polewski, L., Bechtella, L. & Pagel, K. A 3D-printed offline nano-ESI source for bruker MS instruments. *J. Am. Soc. Mass Spectrom.* **34**, 2403–2406 (2023).
53. Ruotolo, B. T., Benesch, J. L. P., Sandercock, A. M., Hyung, S.-J. & Robinson, C. V. Ion mobility–mass spectrometry analysis of large protein complexes. *Nat. Protoc.* **3**, 1139–1152 (2008).
54. Pagel, K. & Harvey, D. J. Ion mobility-mass spectrometry of complex carbohydrates: collision cross sections of sodiated N-linked glycans. *Anal. Chem.* **85**, 5138–5145 (2013).
55. Hofmann, J. et al. Estimating collision cross sections of negatively charged N-glycans using traveling wave ion mobility-mass spectrometry. *Anal. Chem.* **86**, 10789–10795 (2014).
56. Hughes, C. S. et al. Ultrasensitive proteome analysis using paramagnetic bead technology. *Mol. Syst. Biol.* **10**, 757 (2014).
57. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47–e47 (2015).
58. Alboukadel Kassambara. *Rstatix: Pipe-Friendly Framework for Basic Statistical Tests. R Package Version 0.7.2.* <https://rpkgs.datanovia.com/rstatix/> (2023).
59. Varki, A. et al. Symbol nomenclature for graphical representations of glycans. *Glycobiology* **25**, 1323–1324 (2015).
60. Nath, S. & Mukherjee, P. MUC1: a multifaceted oncoprotein with a key role in cancer progression. *Trends Mol. Med.* **20**, 332–42 (2014).
61. Brayman, M., Thathiah, A. & Carson, D. D. MUC1: a multifunctional cell surface component of reproductive tissue epithelia. *Reprod. Biol. Endocrinol.* **2**, 4 (2004).
62. Ponnusamy, M. P., Seshacharyulu, P., Vaz, A., Dey, P. & Batra, S. K. MUC4 stabilizes HER2 expression and maintains the cancer stem cell population in ovarian cancer cells. *J. Ovarian Res.* **4**, 7 (2011).
63. Tytgat, K. M., Opdam, F. J., Einerhand, A. W., Buller, H. A. & Dekker, J. MUC2 is the prominent colonic mucin expressed in ulcerative colitis. *Gut* **38**, 554–563 (1996).

64. Tytgat, K. M. A. J., van der Wal, J.-W. G., Einerhand, A. W. C., Büller, H. A. & Dekker, J. Quantitative analysis of MUC2 synthesis in ulcerative colitis. *Biochem. Biophys. Res. Commun.* **224**, 397–405 (1996).
65. Goodell, C. A. et al. Characterization of the tumor marker muc16 (ca125) expressed by murine ovarian tumor cell lines and identification of a panel of cross-reactive monoclonal antibodies. *J. Ovarian Res* **2**, 8 (2009).
66. Govindarajan, B. et al. A metalloproteinase secreted by *Streptococcus pneumoniae* removes membrane mucin MUC16 from the epithelial glycocalyx barrier. *PLoS ONE* **7**, e32418 (2012).
67. Karlsson, N. G. & Thomsson, K. A. Salivary MUC7 is a major carrier of blood group I type O-linked oligosaccharides serving as the scaffold for sialyl Lewis x. *Glycobiology* **19**, 288–300 (2009).
68. Takehara, S., Yanagishita, M., Podyma-Inoue, K. A. & Kawaguchi, Y. Degradation of MUC7 and MUC5B in Human Saliva. *PLoS ONE* **8**, e69059 (2013).
69. Sheehan, J. K. et al. Identification of molecular intermediates in the assembly pathway of the MUC5AC mucin. *J. Biol. Chem.* **279**, 15698–15705 (2004).
70. Ruchaud-Sparagano, M.-H., Westley, B. R. & May, F. E. B. The trefoil protein TFF1 is bound to MUC5AC in humangastric mucosa. *Cell. Mol. Life Sci.* **61**, 1946–1954 (2004).

Acknowledgements

We thank the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) for financing this work under: the CRC 1449 “Dynamic Hydrogels at Biointerfaces” – Project ID 431232613: L.B., K.F., M.S., S.Y.G., M.A.M., P.M. and K.P. –, the CRC 1340 “Matrix in Vision” – Project ID 372486779: Ł.P. and K.P.–, and G.M.V. and K.P. are grateful for the funding by the European Union’s Horizon 2020 Research and Innovation Programme grant number 899687—HS-SEQ.

Author contributions

O-glycomics: K.P. directed the research. L.B. and K.P. designed the research. L.B., N.G.K. and K.P. supervised the study. L.B., G.R.E. and M.S. conducted the sample preparation. L.B., G.R.E., M.S., Ł.P. and M.G. contributed to the acquisition of experimental data. L.B., C.J. and N.G.K. analyzed data and interpreted the results. L.B., C.J., G.R.E., M.S., Ł.P., M.G., G.M.V., W.B.S., N.G.K. and K.P. contributed to the redaction of the manuscript. Proteomics: P.M. directed and supervised the research. K.F. conducted the sample preparation and the acquisition of experimental data. K.F. analysed data, and L.B. and K.F. interpreted the results. L.B., K.F. and P.M. contributed to the redaction of the manuscript. Sputum

collection: M.A.M. directed the sputum and ethical approval collection. S.Y.G. and M.A.M. supervised and conducted the sputum and ethical approval collection. L.B., S.Y.G. and M.A.M. contributed to the redaction of the manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-46825-4>.

Correspondence and requests for materials should be addressed to Kevin Pagel.

Peer review information *Nature Communications* thanks Gabe Nagy, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024