



Identification of *N*-glycans with GalNAc-containing antennae from recombinant HIV trimers by ion mobility and negative ion fragmentation

David J. Harvey^{1,2} · Anna-Janina Behrens^{1,3} · Max Crispin^{1,4} · Weston B. Struwe^{1,5}

Received: 29 March 2021 / Revised: 10 May 2021 / Accepted: 14 June 2021 / Published online: 30 July 2021

© The Author(s) 2021

Abstract

Negative ion collision-induced dissociation (CID) of underivatized *N*-glycans has proved to be a simple, yet powerful method for their structural determination. Recently, we have identified a series of such structures with GalNAc rather than the more common galactose capping the antennae of hybrid and complex glycans. As part of a series of publications describing the negative ion fragmentation of different types of *N*-glycan, this paper describes their CID spectra and estimated nitrogen cross sections recorded by travelling wave ion mobility mass spectrometry (TWIMS). Most of the glycans were derived from the recombinant glycoproteins gp120 and gp41 from the human immunodeficiency virus (HIV), recombinantly derived from human embryonic kidney (HEK 293T) cells. Twenty-six GalNAc-capped hybrid and complex *N*-glycans were identified by a combination of TWIMS, negative ion CID, and exoglycosidase digestions. They were present as the neutral glycans and their sulfated and $\alpha 2 \rightarrow 3$ -linked sialylated analogues. Overall, negative ion fragmentation of glycans generates fingerprints that reveal their structural identity.

Keywords Negative ion fragmentation · Ion mobility · *N*-Glycans · *N*-Acetylgalactosamine · Human immunodeficiency virus

Introduction

Fragmentation by collision-induced dissociation of negative ions from *N*-linked glycans (those attached to proteins at

asparagine in an Asn-Aaa-Ser(Thr) consensus sequence, where Aaa is any amino acid except proline) has been shown to provide a wealth of diagnostic cross-ring fragments. Each individual glycan spectrum is sufficient for determining most structural features without the use of derivatization such as permethylation or reducing end tagging [1]. Specifically, for non-acidic glycans, the method provides the composition of the glycan in terms of its constituent isobaric and isomeric monosaccharide composition, its topology in terms of antenna branching (which type of triantennary glycan is present, for example), the location of fucose residues, and the presence or absence of bisecting GlcNAc. Some of these structural features are difficult to determine by traditional methods such as exoglycosidase sequencing as these enzymatic reactions are often incomplete plus they are often time-consuming and expensive. For acidic glycans, those with sialic acid can be converted to neutral compounds by methyl ester [2] or amide formation and, depending on the method used, the resulting spectra provide information on sialic acid linkage by utilizing their different reactivities to the derivatizing reagents such that the linkage is revealed simply by mass difference [3, 4]. We have been systematically examining the negative ion

Published in the topical collection *Analytical Characterization of Viruses* with guest editor Joseph Zaia.

✉ David J. Harvey
david.harvey@ndm.ox.ac.uk

¹ Oxford Glycobiology Institute, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK

² Target Discovery Institute, Nuffield Department of Medicine, University of Oxford, Roosevelt Drive, Oxford OX3 7FZ, UK

³ Present address: GlycoEra AG, Grabenstrasse 3, 8952 Schlieren, Switzerland

⁴ School of Biological Sciences, Faculty of Natural and Environmental Sciences, University of Southampton, Highfield Campus, Southampton SO17 1BJ, UK

⁵ Chemistry Research Laboratory, Department of Chemistry, University of Oxford, South Parks Road, Oxford OX1 3TA, UK

fragmentation of different types of *N*-glycans [5–11] and in this report present negative ion collision-induced dissociation (CID) spectra and ion mobility properties of glycans containing *N*-acetylgalactosamine (GalNAc) terminating their antennae and their identification in the human immunodeficiency virus (HIV) envelope glycoprotein (Env) BG505 SOSIP.664 trimers expressed in HEK 293T cells.

The soluble, recombinant BG505 SOSIP.664 trimer is the prototype of a class of native-like, Env spike-mimetic immunogens that is now being pursued in various vaccine-development programs [12, 13]. In a paper from this laboratory [14], a quantitative site-specific *N*-glycosylation analysis identified the *N*-glycans attached to twenty of the twenty-six sites of the SOSIP.664 Env trimer. Among the glycans were a number of hybrid and complex species with GalNAc rather than the more normal galactose terminating the antennae. Here, we examine the fragmentation behaviour of these *N*-glycan structures. Although *N*-glycans whose antennae terminate in GalNAc are common in glycoproteins from some sources such as the human pituitary gland [15, 16], tenascin R [17], glycodelin [18], porcine cardiac tissue [19], glycoproteins expressed in HEK 293 (human protein C [20] and, membrane protein CD9P-1 [21]), HEK 293F cells (L-selectin [22], the receptor binding domain of the SARS-CoV-2 spike glycoprotein [23]), and some non-mammalian organisms [24], we have not found them in gp120 or gp41 expressed in HEK 293 cells until now. The relevant transferase has been detected in these cells by Do et al. [25] and exoglycosidase digestion confirmed the presence of GalNAc in the current samples. Understanding the range of glycan structures on recombinant glycoproteins, such as viral glycoproteins, is important in the characterization of candidate immunogens and for the interpretation of the resulting immune response. Here, we probe their fragmentation properties under negative ion CID and report their nitrogen cross sections recorded by travelling wave ion mobility mass spectrometry (TWIMS).

Materials and methods

Materials

Glycans were released from BG505 SOSIP.664 trimers that were transiently expressed in HEK 293T cells (Thermo Fisher Scientific, Waltham, MA) and purified by 2G12-affinity chromatography followed by size-exclusion chromatography, as previously described [26, 27]. The Env trimers were dissociated into their gp120 and gp41 subunits in SDS-PAGE gels in the presence of dithiothreitol (DTT) and stained with Coomassie blue. The gp41 subunit lacks the transmembrane region and cytoplasmic tail and is referred herein as gp41_{ECTO} (i.e. ectodomain). Following destaining, the bands corresponding to gp120 and gp41_{ECTO} were excised and washed

five times alternatingly with acetonitrile and water. The *N*-glycans were released in situ [28] with peptide *N*-glycosidase F (PNGase F, New England Biolabs, Ipswich MA), at 37 °C for 16 h, according to the manufacturer's instructions. The released glycans were extracted from the gel by washing with water and dried in a SpeedVac concentrator. Reference glycans were obtained from Dextra Laboratories (Reading, UK).

Human pituitary follicle-stimulating hormone (hFSH^{24/21}) was obtained from Dr. G. Bousfield (Wichita State University, KS). The *N*-glycans were released with PNGase F from reduced and carboxymethylated of purified hormone (13–30 µg) as described earlier [16] and desialylated by incubation with sialidase from *Arthrobacter ureafaciens*.

Exoglycosidase digestions

Approximately half of the samples were examined directly by ESI-MS, and the remainder were desialylated with the non-specific neuraminidase from *Clostridium perfringens* (New England Bioscience) according to the manufacturer's instructions. The nature of the monosaccharide constituents was confirmed by exoglycosidase digestions as reported earlier, e.g. [29].

Sample preparation for mass spectrometry

Following release from the glycoproteins, all glycan samples were cleaned with a Nafion® 117 membrane as described earlier by Börnsen et al. [30] before examination by mass spectrometry. They were dissolved in a solution of methanol:water (1:1, v:v) containing ammonium phosphate (0.05 M, to maximize formation of $[M+H_2PO_4]^-$ ions, the most common type of ion normally seen from biological samples). Formation of adducts, such as those from phosphate, gives stable ions and prevents extensive in-source fragmentation commonly seen with deprotonated ions. Samples were then centrifuged at 10,000 rpm (9503×g) for 1 min to sediment any particulate matter.

Ion mobility mass spectrometers

Travelling wave ion mobility mass spectrometry experiments were carried out with a Synapt G2Si travelling wave ion mobility mass spectrometer (Waters, Manchester, UK) [31] fitted with a nano-electrospray (nESI) ion source. Thin-wall nanospray capillaries were used for introducing the samples (Waters, Manchester, UK). Ion source conditions were as follows: ESI capillary voltage, 1.0–1.2 kV cone voltage, 100–180 V, ion source temperature 80 °C. The T-wave velocity and peak height voltages were 450 m/s and 40 V respectively unless otherwise specified. Fragmentation was performed after mobility separation in the transfer cell, with argon as the collision gas. The collision cell voltage (60–130 V) was

adjusted according to the parent ion mass to give an even distribution of fragment ions across the mass range. The instrument was externally mass calibrated with dextran oligomers (Glc₂₋₁₃) from *Leuconostoc mesenteroides* and the same material was used to calibrate the drift cell [32, 33]. Data acquisition and processing were carried out using the Waters DriftScope (version 2.8) software and MassLynx™ (version 4.1). Estimated collisional cross sections (CCS), in nitrogen, were determined with the cross section calculation algorithm supplied with DriftScope or, in the case of calculations made from extracted fragment arrival time distributions (ATDs), by the method reported by Thalassinos et al. [34]. In these latter measurements, it was found that the drift times of the ions measured in the MS and CID spectra were slightly different (approx. 0.46 ms) necessitating a correction to be made [35, 36]. The scheme devised by Domon and Costello [37] was used to name the fragment ions. Interpretation of the fragmentation data followed the rules established earlier [5–7, 38, 39].

Spectra from the FSH sample were recorded with a Waters quadrupole time-of-flight (Q TOF) Ultima Global instrument. Samples (about 50 pmoles/μl) in 1:1 (v:v) methanol:water were infused with Proxeon (Proxeon Biosystems, Odense, Denmark) borosilicate capillaries at about 5 μl/min. The ion source was maintained at 120 °C, the infusion emitter potential was 1.1 kV, the cone was at 100 V, and the RF-1 voltage was set at 180 V. Spectra (2 s scans) were acquired with a digitization rate of 4 GHz and accumulated until a satisfactory signal to noise ratio had been obtained. For CID data acquisition, the parent ion was selected at low resolution (about 5 *m/z* mass window) to allow transmission of isotope peaks and fragmented with argon as collision gas at a pressure (recorded on the instrument's pressure gauge) of 0.5 mBar. The voltage on the collision cell was adjusted with mass and charge to give an even distribution of fragment ions across the mass scale. Typical values were 80–120 V. Other voltages were as recommended by the manufacturer. The instrument was externally calibrated with *N*-glycans released from bovine fetuin and are cited to one decimal place. Data acquisition and processing were performed with MassLynx 4.1 as earlier.

Results and discussion

Ion mobility extraction of *N*-glycan spectra from the HIV samples

The ESI spectrum of the glycans released from gp120 gave a spectrum (Fig. 1b), similar to that obtained earlier [41] from a related sample, although the relative proportions of the glycans differed considerably. Ion mobility allowed the singly (Fig. 1c), doubly (Fig. 1d), and triply charged ions (Fig. 1e) to be separated, aiding spectral interpretation. The ESI spectrum of the glycans released from gp41 (Fig. 2b) was weaker

but, again, ion mobility allowed clear spectra of the singly (Fig. 2c), doubly (Fig. 2d), and triply charged ions (Fig. 2e) to be extracted. In these spectra, neutral glycans produced [M+H₂PO₄][−] ions whereas the singly charged sialylated and sulfated glycans were detected as [M-H][−] species. Doubly and triply charged ions were present mainly as [M-H₂]^{2−} and [M-H₃]^{3−} species respectively although some sialylated glycans produced [M-H+Na+(H₂PO₄)]^{2−} ions. All identified GalNAc-containing glycans, with their respective ions, are listed in Table 1. A full list of the glycans present in these samples has been published in the Supplementary Information section of Behrens et al. 2016 [14].

CID spectra

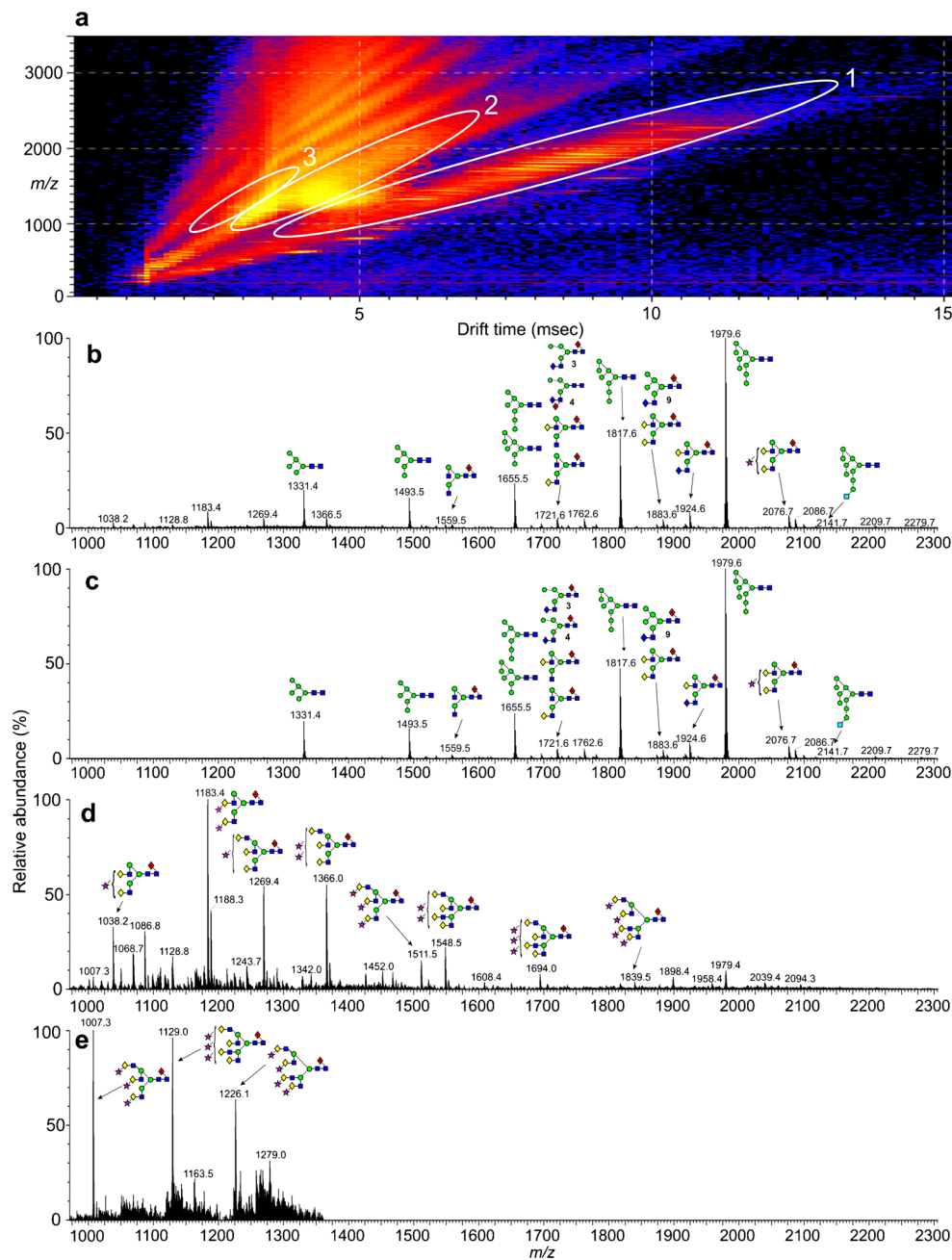
The general features of the negative ion fragmentation of high-mannose and complex *N*-glycans whose antennae terminate in galactose have been reported before [1, 5–7, 39] and will not be repeated here. Suffice it to say that features such as the branching pattern, linkage of sialic acid, the position of fucose substituents, and the presence or absence of bisecting GlcNAc residues are easily determined from the spectra as will be seen from the following discussion.

Complex glycans containing GalNAc

Figure 3 shows a selection of the negative ion CID spectra of several of the glycans whose antennae terminate in GalNAc on both gp120 and gp41. All but that from glycan 1 (panel a) were recorded with the TWIMS instrument whereas that from glycan 1 was obtained with a standard Q-TOF mass spectrometer and, thus, some of the minor ions may be from contaminating compounds. The use of ion mobility to record the other spectra enabled any contaminating ions to be removed.

The major ion identifying the GalNAc-GlcNAc-containing antennae was the ^{1,3}A₃ ion at *m/z* 465.1. This was the major ion in the MS/MS spectrum of glycan 1 (Fig. 3a). The C_{1α} ion at *m/z* 220.1 and the ^{2,4}A₂ ion at *m/z* 262.1 further characterized the structure of the antenna. The core di-*N*-acetylchitobiose region was defined by the ^{2,4}A₆, B₅, and ^{2,4}A₅ ions at *m/z* 1154.3, 1094.3, and 951.3 respectively. This pattern of three ions was repeated in the spectra of all glycans in Fig. 3 and enabled the location of core fucose to be confirmed (see below). Loss of the GalNAc residue from the ^{2,4}A₅ ion to give *m/z* 748.2 further confirmed the presence of GalNAc at the non-reducing terminus of glycan 1. Location of the GalNAc-GlcNAc chain on the mannose residue attached to the 3-position of the core-branching mannose was made by comparison with the spectra of the other glycans, all of which had the GalNAc-GlcNAc located at this position. This conclusion was supported by the absence of a D-18 fragment ion at *m/z* 711, a prominent fragment in the spectra of the glycans containing GalNAc-GlcNAc as the 6-antenna.

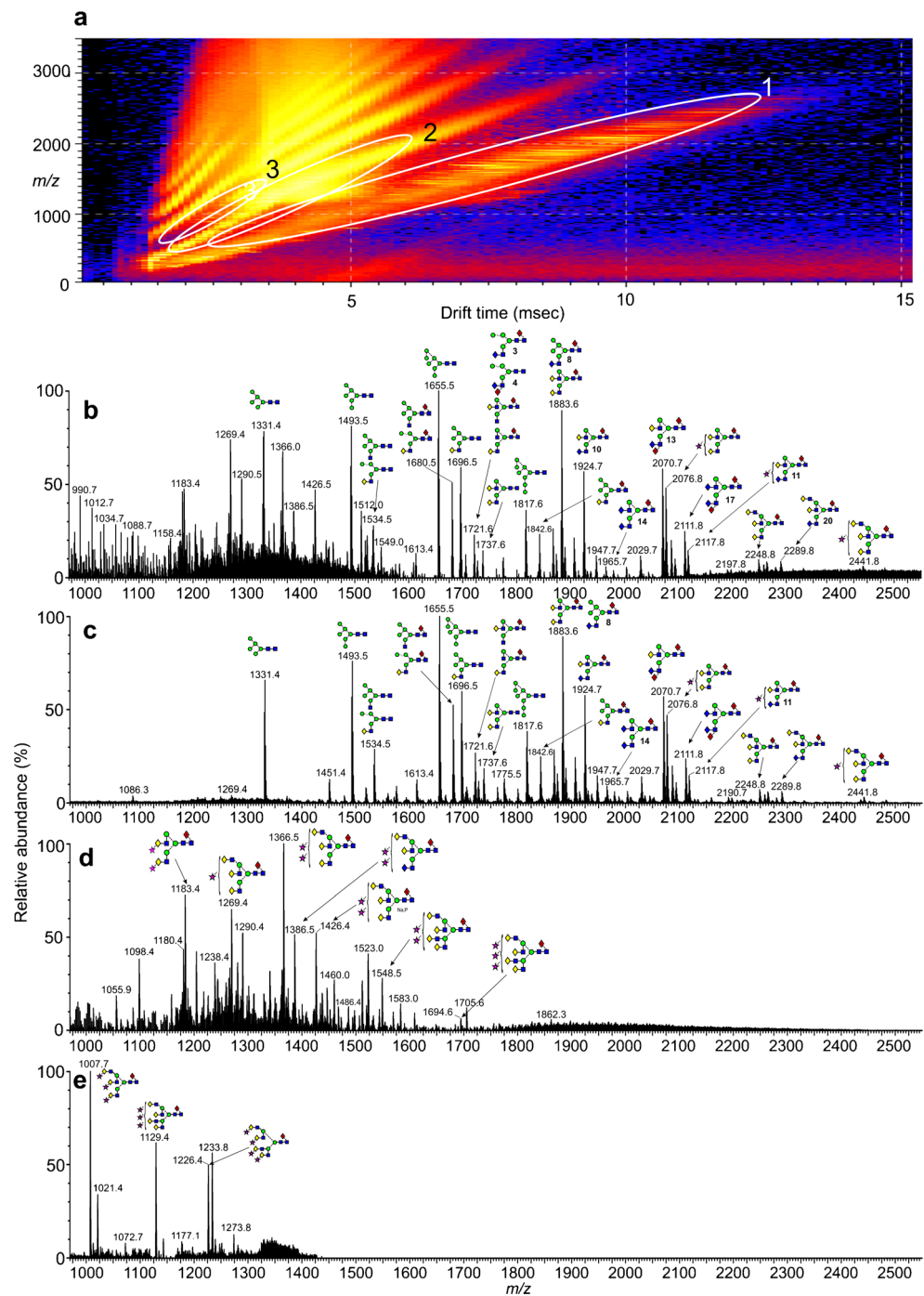
Fig. 1 (a) DriftScope display of *N*-glycans released from gp120. (b) Negative ion ESI spectrum of the *N*-glycans. (c) Mobility-extracted singly charged ions from the spectrum shown in panel b. (d) Mobility-extracted doubly charged ions from the spectrum shown in panel b. (e) Mobility-extracted triply charged ions from the spectrum shown in panel b. Numbers in bold accompanying the structures are those of the glycans listed in Table 1. Symbols used for the glycans are ● = mannose, ■ = GlcNAc, ■ = glucose, ◆ = fucose, ◆ = galactose, ◆ = GalNAc, and ★ = Neu5Ac (sialic acid). Solid lines connecting the symbols are β -linkages; broken lines are α -linkages. The angle of the lines shows the linkage position (for more information, see [40])



The m/z 1924.7 ion ($[M+H_2PO_4]^-$) in the glycan profiles was produced by the biantennary glycan **10** containing one GalNAc residue. Its CID spectrum is shown in Fig. 3b. Two $^{1,3}A_3$ ions were present; the first at m/z 424.1 contained Gal-GlcNAc ($^{1,3}A_{3\beta}$) and the second at m/z 465.1 contained the GalNAc-GlcNAc chain ($^{1,3}A_{3\alpha}$). Two C_1 fragments at m/z 179.1 ($C_{1\beta}$) and 220.1 ($C_{1\alpha}$) were consistent with the two antennae terminating with galactose and GalNAc respectively. Ions generated from the loss of both galactose (m/z 1154.4) and GalNAc (m/z 1113.4) respectively from the $^{2,4}A_5$ ion confirmed the terminal sugar residues at the non-reducing terminus. A common feature of the fragmentation of negative ions from these glycans is the

formation of an ion, termed ion D, formed by formal loss of the *N,N*-di-acetyl-chitobiose (GlcNAc-GlcNAc) core and the antenna linked to the 3-position of the branching mannose. It is accompanied by a second ion formed by loss of H_2O and termed D-18 and two cross-ring fragments of the branching mannose, $^{0,3}A$ and $^{0,4}A$, the latter ions being particularly prominent in the spectra of the high-mannose glycans. The masses of the D and D-18 ions in the spectrum of the biantennary glycan, **10**, were m/z 688.2 and 670.2 respectively showing that the Gal-GlcNAc antenna was located at the 6-position of the branching mannose leaving the GalNAc-GlcNAc antenna on position 3. Location of the GalNAc residue specifically to the 3-antenna of a biantennary

Fig. 2 (a) DriftScope display of *N*-glycans released from gp41. (b) Negative ion ESI spectrum of *N*-glycans released from gp41. (c) Mobility-extracted singly charged ions from the spectrum shown in panel b. (d) Mobility-extracted doubly charged ions from the spectrum shown in panel b. (e) Mobility-extracted triply charged ions from the spectrum shown in panel b. Numbers in bold accompanying the structures are those of the glycans listed in Table 1



glycan has previously been reported from human pituitary follicle-stimulating hormone [15]. The presence of GalNAc in the 6-antenna would have caused a shift of the D and D-18 ions by 41 units to m/z 728.2 and 711.2; no such ions were present. Location of the fucose residue to the 6-position of the reducing terminal GlcNAc residue was revealed by the masses of the $^{2,4}A_6$, B_5 , and $^{2,4}A_5$ ions (m/z 1519.5, 1459.5, and 1316.5 respectively), all of which lacked the fucose residue. In other respects, the spectrum paralleled that of a reference sample of the biantennary glycan with two galactose residues (27). The CID

spectrum of the biantennary glycan 7 lacking the fucose residue, from the FSH sample, was virtually identical.

A fucosylated biantennary glycan ($\text{Hex}_3\text{HexNAc}_6\text{Fuc}_1$) with both antennae terminating in GalNAc (14, m/z 1965.7) was also present in the gp41 and FSH samples. Its CID spectrum (Fig. 3c) exhibited no ion at m/z 424.1 that would be expected from a glycan with Gal-GlcNAc-antennae; instead, this ion was replaced by the prominent $^{1,3}A_3$ ion at m/z 465.1. Absence of the C_1 ion at m/z 179.1 and the presence of m/z 220.1 confirmed that both antennae terminated with GalNAc.

Table 1 Structures and properties of the GalNAc-containing *N*-glycans. *N*, *N*-acetylhexosamine (i.e. HexNAc); *H*, hexose; *F*, fucose; *S*, *N*-acetylneuraminic acid (i.e. Neu5Ac); *SO*, sulfate. Shown below the

composition are the monoisotopic molecular weight, the detected ion type in parentheses, and the estimated nitrogen cross section in bold where the ion abundance was sufficiently high to give a Gaussian ATD

1	2	3	4	5	6
H ₃ N ₄ 1413.5 (a)	H ₄ N ₄ 1575.5 (a)	H ₄ N ₄ F ₁ 1721.6 (a)	H ₄ N ₄ F ₁ 1721.6 (a)	H ₄ N ₄ F ₂ 1867.6 (a) 419.2	H ₅ N ₄ 1737.6 (a) 391.3
7	8	9	10	11	12
H ₅ N ₄ 1737.6 (a) 393.8	H ₅ N ₄ F ₁ 1883.6 (a) 393.8	H ₅ N ₄ F ₁ 1883.6 (a)	H ₄ N ₅ F ₁ 1924.6 (a)	H ₄ N ₅ F ₁ S ₁ 2117.8 (b)	H ₄ N ₅ F ₁ SO ₁ 1906.6 (b) 426.4
13	14	15	16	17	18
H ₄ N ₅ F ₂ 2070.7 (a) 450.0	H ₃ N ₆ F ₁ 1965.7 (a) 437.7	H ₃ N ₆ F ₁ S ₁ 2158.8 (b) 470.1	H ₃ N ₆ F ₁ S ₁ 1947.7 (b) 434.2	H ₃ N ₆ F ₂ 2111.7 (a) 457.8	H ₃ N ₆ F ₂ SO ₁ 2093.7 (b) 451.5
19	20	21	22	23	24
H ₃ N ₆ F ₂ 2257.8 (a) 475.2	H ₅ N ₆ F ₁ 2289.8 (a) 472.1	H ₅ N ₆ F ₁ S ₁ 2482.8 (b)	H ₅ N ₆ F ₁ S ₂ 1386.5 (c)	H ₅ N ₆ F ₁ S ₃ 1021.0 (d)	H ₅ N ₆ F ₁ SO ₁ 2271.8 (b)
	25	26	27	28	
	H ₅ N ₆ F ₂ 2435.8 (a)	H ₅ N ₆ F ₂ S ₁ 2629.0 (b)	H ₅ N ₄ 1737.6 (a)	H ₅ N ₄ F ₁ 1883.6 (a) 416.4	

(a) [M+H₂PO₄]⁻, (b) [M-H]⁻, (c) [M-H₂]²⁻, (d) [M-H₃]³⁻

The C₂ ion appeared at *m/z* 423.1 confirming the presence of the GalNAc-GlcNAc moiety. D and D-18 ions at *m/z* 728.2 and 711.2 confirmed inclusion of GalNAc in the 6-antenna. Jin et al. [19] have reported the spectrum of the [M-2H]²⁻ ion from the corresponding glycan lacking fucose and have also observed the ^{2,4}A₂ (*m/z* 262.1), ^{0,2}A₂ (*m/z* 321.1), C₂ (*m/z* 423.1), and ^{1,3}A₃ (*m/z* 465.1) ions. The singly charged ^{2,4}A₅ ion at *m/z* 1356.4 was of low abundance but the doubly charged version of this ion at *m/z* 678.3 was the base peak. The ^{2,4}A₆ ion, if present, (not labelled) was of very low abundance or absent.

Biantennary glycans were also present with a fucose residue on the antennae, additional to that on the reducing terminal GlcNAc residue. Thus, the spectrum of the difucosylated glycan **13** (*m/z* 2070.7) containing one GalNAc residue is shown in Fig. 3d. The D and D-18 ions appeared at *m/z* 688.2 and 670.2 confirming that the 6-antenna terminated in galactose and that the GalNAc was in the 3-position as in the glycans whose spectra have already been discussed. The presence of fucose on the GalNAc-containing antenna caused a shift of the GalNAc-containing ^{1,3}A₃ ion from *m/z* 465.1 to *m/z* 611.2. No corresponding shift was noted for the galactose-containing ion at *m/z*

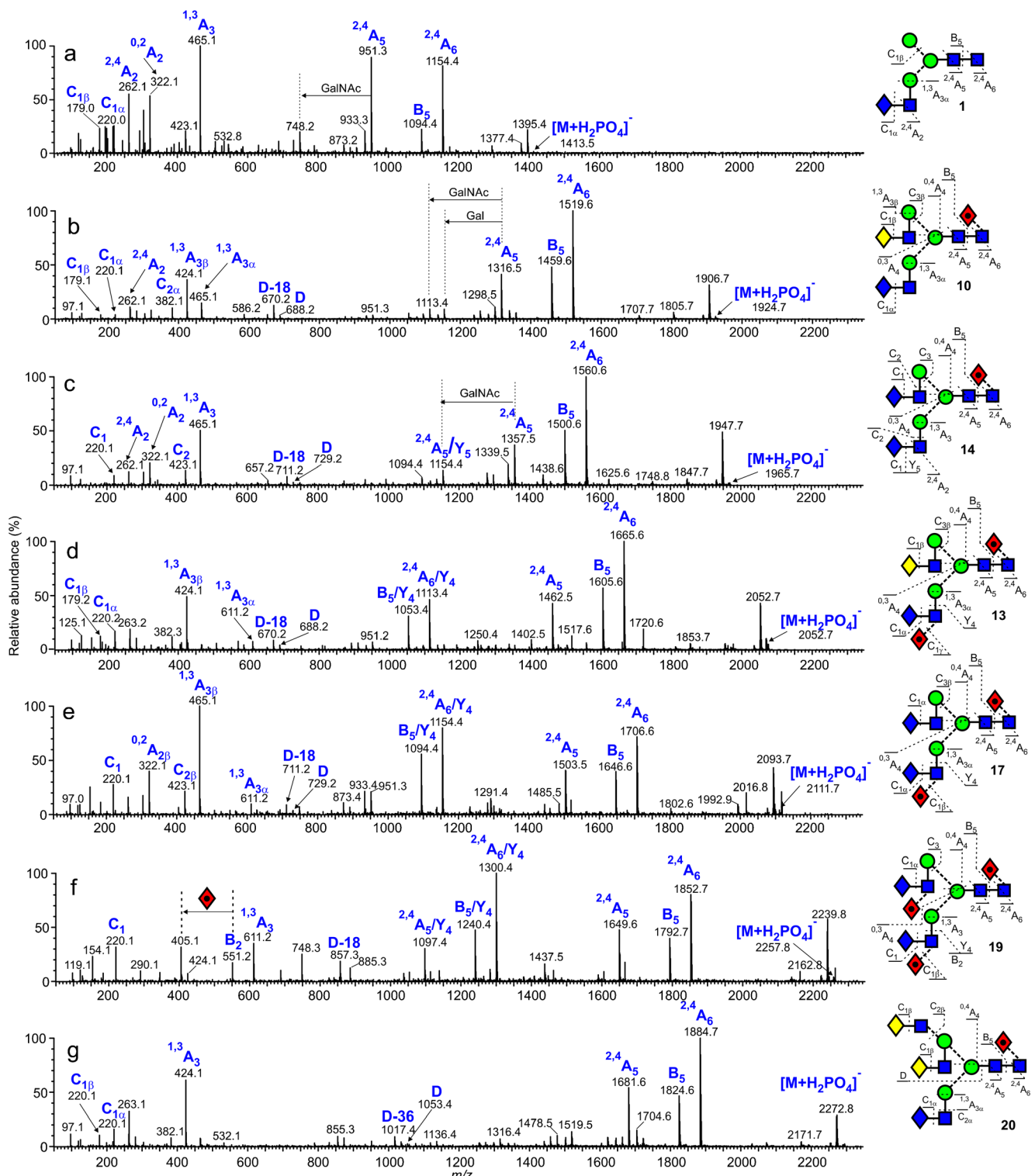


Fig. 3 Negative ion CID spectra of (a) glycan **1** from hFSH, (b) the biantennary glycan **10** from gp41, (c) the biantennary glycan **14** with two GalNAc-capped antennae from gp41, (d) the difucosylated biantennary glycan **17** from gp41, (e) the difucosylated biantennary

glycan **17** from gp41, (f) the trifucosylated biantennary glycan **19** from gp41, and (g) the triantennary glycan **20** from gp41. Symbols for the glycans are as described in the legend to Fig. 1

424.1 showing the absence of fucose on the galactose-containing antenna. The $C_{1\alpha}$ ion at m/z 220.1 showed that the fucose was not located on the terminal GalNAc residue. Work

by other investigators on the structure of *N*-glycans on glycoproteins expressed in HEK 293 cells has also shown that fucose that is attached to an antennae is linked at the 3-position of the

GlcNAc residue [20–22]. Furthermore, fucose substitution on galactose has been found to produce a prominent pair of ions at m/z 427.1 and 409.1 produced by an $^{0,2}A_3$ cleavage followed by loss of H_2O [42]. These ions and the corresponding predicted ones (m/z 468.1 and 450.1) from a GalNAc-containing antenna were absent, leaving little doubt that the attachment of fucose was to the 3-position of the GlcNAc residue. We have found fucose at this position in glycans from the human parotid gland (Gal-GlcNAc-antennae) [7] and in their spectra the abundance of the ions formed by loss of the fucosylated 3-antenna (GalNAc-(Fuc)GlcNAc) from the $^{2,4}A_6$, B_5 , and $^{2,4}A_5$ ions is considerably elevated compared with the spectra of corresponding compounds where there is no fucose attached to the GlcNAc residue [7]. The same phenomenon was noted in these GalNAc-containing spectra (m/z 1113.4, 1053.4, and 951.3 in spectrum 3d).

Glycan **17** (m/z 2111.7) with two GalNAc and two fucose residues gave a spectrum (Fig. 3e) with similar characteristics again indicating fucose substitution at the 3-position of the GlcNAc of the 3-antenna. The D and D-18 ions appeared at m/z 729.2 and 711.2 confirming that the 6-antenna terminated in GalNAc and showing no fucose substitution on this antenna. Again, loss of GalNAc-(Fuc)GlcNAc from the $^{2,4}A_6$, B_5 , and $^{2,4}A_5$ ions produced major fragments at m/z 1154.4, 1094.4, and 951.3 respectively.

A fucosylated GalNAc-containing glycan (HexNAc₆Hex₃Fuc₃) containing three fucose residues gave the $[M+H_2PO_4]^-$ ion at m/z 2257.8. Its CID spectrum (Fig. 3f) contained no $^{1,3}A_3$ ion at m/z 465.1 (unsubstituted antenna) but a prominent one at m/z 611.2 indicating fucose linked to an antenna. No di-fucosylated equivalent (m/z 757.3) was detected showing that each antenna contained a single fucose residue. Again the C_1 ion at m/z 220.1 (no fucose) and the absence of the $^{0,2}A_3$ ions at m/z 468.2 and 450.2 showed that the fucose residues were located on the GlcNAc residues to give structure **19**.

The $[M+H_2PO_4]^-$ ion at m/z 2289.8 resolved to a composition of Hex₅HexNAc₆Fuc₁. Its CID spectrum (Fig. 3g) contained $^{1,3}A_3$ ions at m/z 424.1 and 465.1 showing the presence of Gal-GlcNAc and GalNAc-GlcNAc-antenna. D, D-18, and D-36 ions at m/z 1053.4, 1035.4, and 1017.4 respectively were consistent with a triantennary glycan containing an unsubstituted branched 6-antenna (two Gal-GlcNAc chains) [43] leaving GalNAc-GlcNAc as the structure of the 3-antenna as found in the biantennary glycan **10**. Location of the fucose residue to the core was indicated by the masses of the $^{2,4}A_6$, B_5 , and $^{2,4}A_5$ ions at m/z 1852.7, 1792.7, and 1649.6 respectively. This glycan was, therefore, identified as **20**. The weak CID spectrum (not shown) of a very minor glycan containing two fucose residues (m/z 2435.8, (HexNAc₆Hex₃Fuc₂)) contained D-36 (m/z 1017.4) and $^{1,3}A_3$ (m/z 611.2) ions supporting the structure **25**.

Small amounts of different hybrid glycans containing GalNAc were also detected from both gp41 and gp120, but

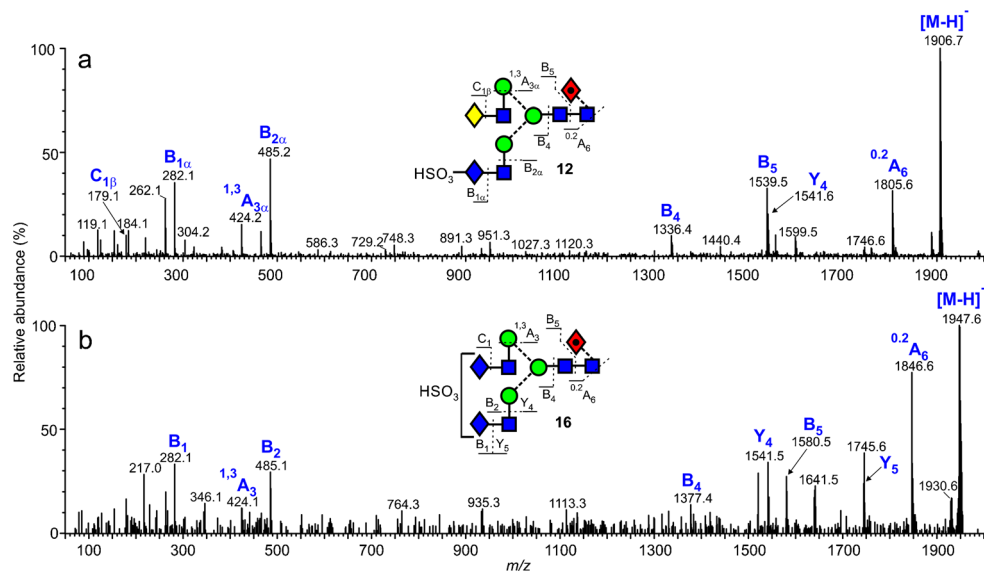
these were isobaric with other major glycans and individual, clean spectra were not obtained. However, by comparing the spectra with those of reference spectra (Dextra Laboratories) of the major contaminating glycans, the ions belonging to the hybrid glycans could be identified. In some cases, there was a difference in drift times allowing the relevant fragment ions to be isolated. Thus, the hybrid glycan **6** (HexNAc₄Hex₅, m/z 1737.6), isobaric with the biantennary glycan **7**, was identified by the presence of the usual D, D-18, $^{0,3}A_3$, and $^{0,4}A_3$ ions (m/z 647.2, 629.2, 575.2, and 545.2 respectively) showing a mannose-containing 6-antenna as in the CID spectrum of the high-mannose glycan, Man₅GlcNAc₂, which has the same mannose residues in the 6-antenna [6] and the GalNAc-GlcNAc-containing $^{1,3}A_3$ ion at m/z 465.1. The sets of fragment ions from the two compounds showed slight separation. The cross section measured in the MS spectrum for the mixture of glycans **6** and **7** was 392.2 \AA^2 . Extracted fragment ATDs of the $^{1,3}A_3$ ions (m/z 424.1 for **7** and 465.2 for **6**) in the CID spectrum, diagnostic for the two compounds, allowed further measurements of the cross sections of these two compounds to be made. These measurements gave values of 393.8 \AA^2 for the biantennary glycan **7** and 391.3 \AA^2 for the hybrid glycan **6**.

The fucosylated analogues of compound **6** (m/z 1883.6) are isobaric with several other potential glycans including the fucosylated biantennary glycan **28**. In addition to the ions characterizing this latter glycan in Fig. 3a, the spectrum contained very minor D, D-18, $^{0,3}A_3$, and $^{0,4}A_3$ ions at m/z 647.2, 629.2, 575.2, and 545.2 characterizing the mannose residues of the 6-antenna of compounds such as **8** and **9**. This conclusion was consistent with the presence of minor $^{2,4}A_6$, B_5 , and $^{2,4}A_5$ ions at m/z 1624.6, 1564.6, and 1421.6 respectively whose ion mobility drift times aligned with those of m/z 611.2 but differed from those of the glycans with a fucosylated core GlcNAc. These observations provided evidence for the presence of glycan **9**. Although the core-fucosylated glycan **8** was probably present, its presence could not be confirmed. The $^{1,3}A_3$ ion at m/z 465.1 diagnostic for GalNAc-GlcNAc was of very low relative abundance and the $^{2,4}A_6$, B_5 , and $^{2,4}A_5$ ions were isobaric with those from the biantennary glycan **28** and showed no mobility separation.

Analogues of all of these GalNAc-containing hybrid glycans containing one fewer mannose residue in the 6-antenna and with zero (**2**, m/z 1575.5), one (**3** and **4**, m/z 1721.6, Fig. 3b), and two (**5**, m/z 1867.6) fucose residues were detected by the presence of the appropriate diagnostic fragment ions as discussed earlier.

Sulfated glycans

Released *N*-glycans whose antennae terminate in GalNAc often contain a 4-linked sulfate residue [17, 22, 44]. Several

Fig. 4 Negative ion CID spectra of the sulfated glycans **12** and **16**

such glycans were found in the mixture of glycans from gp120 and gp41. Thus, Fig. 4a shows the spectrum of the glycan **12** (m/z 1906.6, $[M-H]^-$ ion) of composition HexNAc₅Hex₄Fuc₁-SO₃H. The spectrum contained prominent fragment ions at m/z 282.1 ($B_{1\alpha}$, HSO₃-GalNAc) and 485.2 ($B_{2\alpha}$, HSO₃-GalNAc-GlcNAc) as found earlier for sulfated glycans of this type [9]. Ions at the high-mass end of the spectrum included the $^{0,2}A_6$ ion at m/z 1805.6, an ion frequently present in the negative ion spectra of acidic glycans. Other ions were mainly B fragments with the B_5 fragment consistent with the presence of fucose on the core GlcNAc. Glycan **16** had a similar structure but with an additional GalNAc residue in place of the galactose. Its spectrum (Fig. 4b) was similar to that of glycan **12**. Additional Y ions at m/z 1745.6 and 1541.6 arose from loss of GalNAc and GalNAc-GlcNAc respectively from the antenna not containing the sulfate group. From the rather weak spectrum, it was not possible to determine which antenna carried the sulfate group. The sulfated glycans **18** (m/z 2093.7) and **24** (m/z 2271.8) were identified similarly.

Sialylated glycans

Most of the above neutral glycans produced sialylated analogues. All were sensitive to sialidase, giving the corresponding neutral glycans. Monosialylated glycans (**11**, **15**, **21**, and **26**) gave $[M-H]^-$ ions. Doubly and triply charged negative deprotonated ions were formed by the more highly sialylated glycans. Both identified compounds (**22** and **23**) were sialylated derivatives of the neutral glycans discussed above. The CID spectra of all sialylated glycans contained a prominent B_1 fragment at m/z 290.1 defining the Neu5Ac group and the absence of a fragment ion at m/z 306.2 in the samples from the HIV glycans showed that all sialic acid groups were $\alpha 2 \rightarrow$

3-linked [9, 45]. $\alpha 2 \rightarrow 6$ -linked sialic acid was detected in the FSH samples [16].

Conclusions

Gp120 and gp41 from BG505 SOSIP.664 trimers produced in HEK 293T cells were found to contain a complex mixture of *N*-glycans including 25 hybrid and complex glycans whose antennae terminated in GalNAc. They were easily identified by the presence of an abundant $^{1,3}A$ cross-ring fragment at m/z 465.1. Fucose substitution on an antenna produced a shift of this ion to m/z 611.2. In other respects, these compounds fragmented in a similar manner to other hybrid and complex glycans reported earlier. D and D-18 ions, diagnostic of the composition of the 6-antenna, were present and could be used to define the branching pattern of the triantennary glycans. The masses of these ions showed that the location of the GalNAc residue in glycans with one GalNAc was invariably on the 3-antenna, a property that would be difficult to determine by traditional positive ion fragmentation methods. However, recent methods such as electronic excitation dissociation (EED) do appear to produce ions that allow this structural feature to be determined directly [46]. The current paper, therefore, provides a further illustration of how ion mobility and negative ion CID, combined with exoglycosidase digestions to determine the nature of the monosaccharide constituents, can be used to provide detailed structures of these compounds without the need for extensive fractionation or derivatization.

Abbreviations ATD, Arrival time distribution; CCS, Collisional cross section; CID, Collision-induced dissociation; EED, Electronic excitation

dissociation; *ESI*, Electrospray; *Fuc*, Fucose; *Gal*, Galactose; *GalNAc*, *N*-Acetylgalactosamine; *Glc*, Glucose; *GlcNAc*, *N*-Acetylglucosamine; *HEK*, Human embryonic kidney; *Hex*, Hexose; *HexNAc*, *N*-Acetylhexosamine; *hFSH*, Human pituitary follicle-stimulating hormone; *HIV*, Human immunodeficiency virus; *MS*, Mass spectrometry; *Neu5Ac*, *N*-Acetylneuraminic acid (sialic acid); *PNGase F*, Peptide *N*-glycosidase F; *Q*, Quadrupole; *RF*, Radio frequency; *TOF*, Time-of-flight; *TWIMS*, Travelling wave ion mobility mass spectrometry

Acknowledgements We thank Dr. G. Bousfield (Wichita State University, KS) for permission to use the spectrum of Man₃GlcNAc₃GalNAc₁.

Availability of data and material Not applicable.

Code availability Not applicable.

Funding The work was supported by a the International AIDS Vaccine Initiative (IAVI) through grant INV-008352/OPP1153692 and the IAVI Neutralizing Antibody Center through the Collaboration for AIDS Vaccine Discovery grant OPP1196345/INV-008813, both funded by the Bill and Melinda Gates Foundation. This work was also supported by the National Institute for Allergy and Infectious Diseases through the Scripps Consortium for HIV Vaccine Development (CHAVD) (AI144462). A.J.B is a recipient of a Chris Scanlan Memorial Scholarship from Corpus Christi College, Oxford.

Declarations

Conflict of interest The authors declare no competing interests.

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/>.

References

- Harvey DJ. Negative ion mass spectrometry for the analysis of *N*-linked glycans. *Mass Spectrom Rev.* 2020;39:586–679.
- Powell AK, Harvey DJ. Stabilisation of sialic acids in *N*-linked oligosaccharides and gangliosides for analysis by positive ion matrix-assisted laser desorption-ionization mass spectrometry. *Rapid Commun Mass Spectrom.* 1996;10:1027–32.
- Wheeler SF, Domann P, Harvey DJ. Derivatization of sialic acids for stabilization in matrix-assisted laser desorption/ionization mass spectrometry and concomitant differentiation of $\alpha(2-3)$ and $\alpha(2-6)$ isomers. *Rapid Commun Mass Spectrom.* 2009;23:303–12.
- de Haan N, Yang S, Cipollo J, Wührer M. Glycomics studies using sialic acid derivatization and mass spectrometry. *Nat Rev.* 2020;4:229–42.
- Harvey DJ. Fragmentation of negative ions from carbohydrates: Part 2, Fragmentation of high-mannose *N*-linked glycans. *J Am Soc Mass Spectrom.* 2005;16:631–46.
- Harvey DJ. 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.* 2005;16:622–30.
- Harvey DJ. Fragmentation of negative ions from carbohydrates: Part 3, Fragmentation of hybrid and complex *N*-linked glycans. *J Am Soc Mass Spectrom.* 2005;16:647–59.
- Harvey DJ, Jaeken J, Butler M, Armitage AJ, Rudd PM, Dwek RA. Fragmentation of negative ions from *N*-linked carbohydrates, Part 4. Fragmentation of complex glycans lacking substitution on the 6-antenna. *J Mass Spectrom.* 2010;45:528–35.
- Harvey DJ, Rudd PM. Fragmentation of negative ions from *N*-linked carbohydrates. Part 5: Anionic *N*-linked glycans. *Int J Mass Spectrom.* 2011;305:120–30.
- Harvey DJ, Edgeworth M, Krishna BA, Bonomelli C, Allman S, Crispin M, et al. Fragmentation of negative ions from *N*-linked carbohydrates: Part 6: Glycans containing one *N*-acetylglucosamine in the core. *Rapid Commun Mass Spectrom.* 2014;28:2008–18.
- Harvey DJ, Abrahams JL. Fragmentation and ion mobility properties of negative ions from *N*-linked carbohydrates: Part 7: Reduced glycans. *Rapid Commun Mass Spectrom.* 2016;30:627–34.
- Sanders RW, Derking R, Cupo A, Julien JP, Yasmeen A, de Val N, et al. A next-generation cleaved, soluble HIV-1 Env trimer, BG505 SOSIP.664 gp140, expresses multiple epitopes for broadly neutralizing but not non-neutralizing antibodies. *PLoS Pathog.* 2013;9:e1003618.
- Sanders RW, van Gils MJ, Derking R, Sok D, Ketas TJ, Burger JA, et al. HIV-1 neutralizing antibodies induced by native-like envelope trimers. *Science.* 2015;349:aac4223.
- Behrens A-J, Vasiljevic S, Pritchard LK, Harvey DJ, Andev RS, Krumm SA, et al. Composition and antigenic effects of individual glycan sites of the trimeric HIV-1 envelope. *Cell Rep.* 2016;14:2695–706.
- Butnev VY, Butnev VY, May JV, Shuai B, Tran P, White WK, et al. Production, purification, and characterization of recombinant hFSH glycoforms for functional studies. *Molec Cell Endocrin.* 2015;405:42–51.
- Bousfield GR, Butnev VY, White WK, Smalter Hall A, Harvey DJ. Comparison of follicle-stimulating hormone glycosylation microheterogeneity by quantitative negative mode nano-electrospray mass spectrometry of peptide-*N*-glycanase-released oligosaccharides. *J Glycom Lipidom.* 2015;5:Article 1000129.
- Woodworth A, Fiete D, Baenziger JU. Spatial and temporal regulation of tenascin-R glycosylation in the cerebellum. *J Biol Chem.* 2002;277:50941–7.
- Dell A, Morris HR, Easton RL, Panico M, Patankar M, Oehninger S, et al. Structural analysis of the oligosaccharides derived from glycodefin, a human glycoprotein with potent immunosuppressive and contraceptive activities. *J Biol Chem.* 1995;270:24116–26.
- Jin C, Cherian RM, Liu J, Playà-Albinyana H, Galli C, Karlsson NG, et al. Identification by mass spectrometry and immunoblotting of xenogeneic antigens in the *N*- and *O*-glycomes of porcine, bovine and equine heart tissues. *Glycoconj J.* 2020;37:485–98.
- Yan SB, Chao YB, van Halbeek H. Novel Asn-linked oligosaccharides terminating in GalNAc β (1-4)[Fuc α (1-3)]GlcNAc β (1-) are present in recombinant human protein C expressed in human kidney 293 cells. *Glycobiology.* 1993;3:597–608.
- André M, Morelle W, Planchon S, Milhiet P-E, Rubinstein E, Mollicone R, et al. Glycosylation status of the membrane protein CD9P-1. *Proteomics.* 2007;7:3880–95.
- Wedepohl S, Kaup M, Riese SB, Berger M, Demedde J, Tauber R, et al. *N*-Glycan analysis of recombinant L-selectin reveals sulfated GalNAc and GalNAc-GalNAc motifs. *J Proteome Res.* 2010;9:3403–11.
- Lenza MP, Oyenarte I, Diercks T, Quintana JI, Gimeno A, Coelho H, et al. Structural characterization of *N*-linked glycans in the receptor binding domain of the SARS-CoV-2 spike protein and their

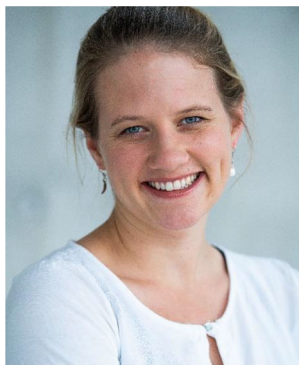
- interactions with human lectins. *Angew Chem Int Ed.* 2020;59:23763–71.
24. Nyame K, Smith DF, Damian RT, Cummings RD. Complex-type asparagine-linked oligosaccharides in glycoproteins synthesized by *Schistosoma mansoni* adult males contain terminal beta-linked *N*-acetylgalactosamine. *J Biol Chem.* 1989;264:3235–43.
 25. Do K-Y, Do S-I, Cummings RD. Differential expression of LacdiNAc sequences (GalNAc β 1-4GlcNAc-R) in glycoproteins synthesized by Chinese hamster ovary and human 293 cells. *Glycobiology.* 1997;7:183–94.
 26. Chung NPY, Matthews K, Kim HJ, Ketas TJ, Golabek M, de Los RK, et al. Stable 293 T and CHO cell lines expressing cleaved, stable HIV-1 envelope glycoprotein trimers for structural and vaccine studies. *Retrovirology.* 2014;11:Article 33.
 27. Dey AK, Cupo A, Ozorowski G, Sharma VK, Behrens AJ, Go EP, et al. cGMP production and analysis of BG505 SOSIP.664, an extensively glycosylated, trimeric HIV-1 envelope glycoprotein vaccine candidate. *Biotechnol Bioeng.* 2018;115:885–99.
 28. Küster B, Wheeler SF, Hunter AP, Dwek RA, Harvey DJ. Sequencing of *N*-linked oligosaccharides directly from protein gels: in-gel deglycosylation followed by matrix-assisted laser desorption/ionization mass spectrometry and normal-phase high performance liquid chromatography. *Anal Biochem.* 1997;250:82–101.
 29. Ritchie G, Harvey DJ, Stroehrer U, Feldmann F, Feldmann H, Wahl-Jensen V, et al. Identification of *N*-glycans from Ebola virus glycoproteins by matrix-assisted laser desorption/ionization time-of-flight and negative ion electrospray tandem mass spectrometry. *Rapid Commun Mass Spectrom.* 2010;24:571–85.
 30. Börmens KO, Mohr MD, Widmer HM. Ion exchange and purification of carbohydrates on a Nafion^(R) membrane as a new sample pretreatment for matrix-assisted laser desorption-ionization mass spectrometry. *Rapid Commun Mass Spectrom.* 1995;9:1031–4.
 31. Giles K, Pringle SD, Worthington KR, Little D, Wildgoose JL, Bateman RH. Applications of a travelling wave-based radio-frequency-only stacked ring ion guide. *Rapid Commun Mass Spectrom.* 2004;18:2401–14.
 32. Hofmann J, Struwe WB, Scarff CA, Scrivens JH, Harvey DJ, Pagel K. Estimating collision cross sections of negatively charged *N*-glycans using travelling wave ion mobility-mass spectrometry. *Anal Chem.* 2014;86:10789–95.
 33. Pagel K, Harvey DJ. Ion mobility mass spectrometry of complex carbohydrates - collision cross sections of sodiated *N*-linked glycans. *Anal Chem.* 2013;85:5138–45.
 34. Thalassinou K, Grabenauer M, Slade SE, Hilton GR, Bowers MT, Scrivens JH. Characterization of phosphorylated peptides using traveling wave-based and drift cell ion mobility mass spectrometry. *Anal Chem.* 2009;81:248–54.
 35. Harper B, Neumann EK, Stow SM, May JC, McLean JA, Solouki T. Determination of ion mobility collision cross sections for unresolved isomeric mixtures using tandem mass spectrometry and chemometric deconvolution. *Anal Chim Acta.* 2016;939:64–72.
 36. Harvey DJ, Scarff CA, Edgeworth M, Struwe WB, Pagel K, Thalassinou K, et al. Travelling-wave ion mobility and negative ion fragmentation of high mannose *N*-glycans. *J Mass Spectrom.* 2016;51:219–35.
 37. Domon B, Costello CE. A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconj J.* 1988;5:397–409.
 38. Bitto D, Harvey DJ, Halldorsson S, Doores KJ, Huisken JT, Bowden TA, et al. Determination of *N*-linked glycosylation in viral glycoproteins by negative ion mass spectrometry and ion mobility. *Methods Molec Biol.* 2015;1331:93–121.
 39. Harvey DJ, Royle L, Radcliffe CM, Rudd PM, Dwek RA. Structural and quantitative analysis of *N*-linked glycans by MALDI and negative ion nanospray mass spectrometry. *Anal Biochem.* 2008;376:44–60.
 40. Harvey DJ, Merry AH, Royle L, Campbell MP, Dwek RA, Rudd PM. Proposal for a standard system for drawing structural diagrams of *N*- and *O*-linked carbohydrates and related compounds. *Proteomics.* 2009;9:3796–801.
 41. Harvey DJ, Crispin M, Bonomelli C, Scrivens JH. Ion mobility mass spectrometry for ion recovery and clean-up of MS and MS/MS spectra obtained from low abundance viral samples. *J Am Soc Mass Spectrom.* 2015;26:1754–67.
 42. Harvey DJ, Struwe WB. Structural studies of fucosylated *N*-glycans by ion mobility mass spectrometry and collision-induced fragmentation of negative ions. *J Am Soc Mass Spectrom.* 2018;29:1179–93.
 43. Harvey DJ, Crispin M, Scanlan C, Singer BB, Lucka L, Chang VT, et al. Differentiation between isomeric triantennary *N*-linked glycans by negative ion tandem mass spectrometry and confirmation of glycans containing galactose attached to the bisecting (β 1-4-GlcNAc) residue in *N*-glycans from IgG. *Rapid Commun Mass Spectrom.* 2008;22:1047–52.
 44. Bousfield GR, Butnev VY, Butnev VY, Nguyen VT, Gray CM, Dias JA, et al. Differential effects of subunit asparagine⁵⁶ oligosaccharide structure on equine lutropin and follitropin hybrid conformation and receptor-binding activity. *Biochemistry.* 2004;43:10817–33.
 45. Wheeler SF, Harvey DJ. Negative ion mass spectrometry of sialylated carbohydrates: discrimination of *N*-acetylneuraminic acid linkages by matrix-assisted laser desorption/ionization-time-of-flight and electrospray-time-of-flight mass spectrometry. *Anal Chem.* 2000;72:5027–39.
 46. Wei J, Tang Y, Ridgeway ME, Park MA, Costello CE, Lin C. Accurate identification of isomeric glycans by trapped ion mobility spectrometry-electronic excitation dissociation tandem mass spectrometry. *Anal Chem.* 2020;92:13211–20.

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



David Harvey obtained his BSc and MSc degrees from the Department of Chemistry, University of Hull, UK, and his PhD degree at the University of Glasgow. He was then appointed MRC Research fellow, Department of Chemistry, University of Glasgow, for 1 year where he worked on GC/MS analysis of steroids. He then moved to Baylor College of Medicine, Houston, TX, continuing work on steroids and expanding into drug metabolism.

He then set up the mass spectrometry laboratory at the Department of Pharmacology, University of Oxford, as Research Officer and then Senior Lecturer where he worked on metabolism and pharmacokinetics of drugs of abuse as studied by GC/MS, and the development of new mass spectrometric methods for lipid analysis. The next move was to the Glycobiology Institute, University of Oxford, as Senior Research Associate and later Reader for FAB and GC/MS studies of *N*-glycans and glycoproteins. This research progressed into MALDI analysis of glycans and later into electrospray and ion mobility studies. Currently, he is a Visiting Professor at the Target Discovery Institute, Nuffield Department of Medicine at Oxford.



Anna-Janina Behrens currently works as a Project Manager at GlycoEra in Schlieren (Switzerland). Starting with her PhD at the University of Oxford and Postdoc at New England Biolabs, she has been working for several years in the fields of glycobiology and mass spectrometry.



Max Crispin is Professor of Glycobiology within the School of Biological Sciences at the University of Southampton, UK, Professor Adjunct at Scripps Research, and a Supernumerary Fellow at Oriel College, Oxford. His research focusses on the development of glycoprotein-based vaccines and biologics.



Weston Struwe is a UKRI Future Leaders Fellow at the University of Oxford, Department of Chemistry, and Stipendiary Lecturer at Brasenose College, Oxford. Previously, he was Chief Scientific Officer of Refeyn, a University of Oxford spin-out based on mass photometry—a single molecule mass imaging technique he helped develop. His research centers on characterizing the glycosylation of viral proteins and identifying how glycans influence interactions with host receptors, innate immune lectin receptors, and anti-viral lectins.

fluency interactions with host receptors, innate immune lectin receptors, and anti-viral lectins.