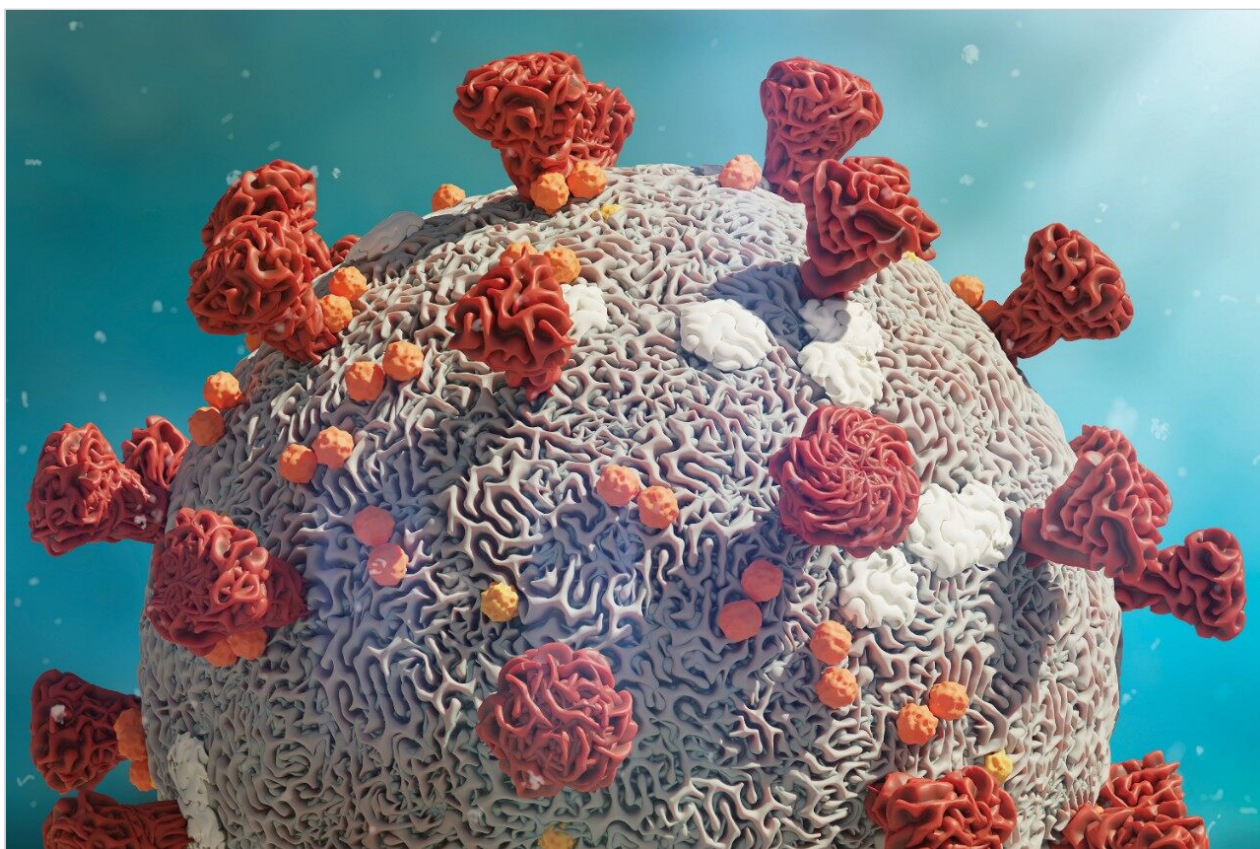Waters™

Comprehending COVID-19: Structural Characterization of the Glycan Assemblies of O-linked Glycopeptides Using Cyclic Ion Mobility

Lindsay Morrison

Waters Corporation, Georgetown University Medical Center, Lombardi Comprehensive Cancer Center

For research use only. Not for use in diagnostic procedures.

Need Help? To learn more about how Waters can help you in your efforts against COVID-19, please contact the COVID-19 Innovation Response Team < https://waterscorp.wufoo.com/forms/z6u1ou20vnme67/>

Abstract

The heavily glycosylated SARS-CoV-2 spike protein is implicated in COVID-19 infectivity and consequently numerous efforts to characterize the glycosylation patterns of the capsid spike protein are being pursued in an effort to guide vaccine and immunotherapy development. Here, the unique configuration of the novel Waters SELECT SERIES Cyclic IMS instrument is exploited for characterization of the glycan structures and linkages of O-linked glycopeptides found near the furin protease cleavage site of the spike protein; furin cleavage exposes the fusion sequence of the spike protein that results in separation of the S1 and S2 subunits is thought to be needed for cell entry. Additionally, extended core 1 and core 2 type structures with both α2-3 and α2-6 sialic acid linkages are identified; glycosylation has been shown to alter substrate recognition and the ratios of the core 2 and extended core 1 structures may impact the virulence of the novel coronavirus.

Benefits

The unique design of the Cyclic IMS system enables site-specific, linkage-specific characterization of the glycan structures of low-level O-linked glycopeptides from the SARS-CoV-2 spike protein. Trap fragmentation is used to release oxonium ion fragments from the O-glycan moiety and high-resolution ion mobility is used to separate and analyze the fragment ions.

Introduction

The outbreak of the recent COVID-19 pandemic, caused by the SARS-CoV-2 coronavirus, has resulted in the sickening of millions and the deaths, to date, of over 2 million people globally.[1] Because of its relative novelty, the scientific community has been engaged in rapid characterization and compilation of the biophysical characteristics of the virus, in hopes of aiding the development of effective vaccine and immunotherapies. The SARS-CoV-2 virion enters the cell by binding to the host cell ACE2 (angiotensin-converting enzyme 2) receptor with a protrusive transmembrane spike protein.[2] The SARS-CoV-2 coronavirus spike protein is a homotrimeric class I fusion protein consisting of two heavily glycosylated subunits, S1 and S2.[3,4] Previous studies of known viral pathogens, including influenza, have revealed the glycan composition of the outer envelope plays a key role in immunoevasion, notably through steric hindrance of immunorecognition sites.[5–7]

The characterization effort of the SARS-CoV-2 pathogen has included several comprehensive studies of the glycosylation of the coronavirus spike protein. To date, these studies have indicated consensus occupancy of 14 of the 22 available N-linked glycosylation sites and conflicting occupancy of the remaining seven sites.[8–16] In addition, Shajahan *et. al* showed evidence for three O-linked glycosylation sites on the S1 domain, one of which was confirmed by Sanda *et. al.*[10,11] In a sequence analysis of the SARS-CoV-spike protein, Andersen and co-workers predicted a furin cleavage site, a site critical to protein activation, in the linker region flanking the S1 and S2 subunits.[2] This region containing the furin cleavage site was further predicted to contain up to three additional O-linked glycosylation sites, the function of which was speculated to be related to infectivity and transmissibility.[2] Characterization of O-linked glycans is notoriously difficult for several reasons: the lack of a consensus sequence, high heterogeneity among the glycoforms, and low relative abundance. Consequently, characterization of these types of analytes requires high performance chromatography and high resolution, high sensitivity mass spectrometry. Previous studies using the Waters SYNAPT platform have shown the capacity to resolve branched and linear glycan structures of isomeric glycans and separate α2-3 and α2-6 linkage isomers from glycan standards.[12–16] MS/MS fragmentation with the Waters SELECT SERIES Cyclic IMS provides clear evidence of O-linked glycosylation along the linker region preceding the polybasic furin cleavage site and cyclic ion mobility is used to site-specifically separate and characterize the O-linked glycan structures. The scalable resolution afforded by this unique experiment provides evidence for a mixture of core 1, extended core 1, and core 2 structures, with resolution of NeuAcα2-3Galβ1-3GalNAc and NeuAcα2-6Galβ1-4GlcNAc, isomers that feature highly similar collisional cross sections.

Experimental

Sample Description

Recombinant SARS-CoV-2 spike protein was expressed in HEK 293 cells and obtained from Acrobiosystems. The disulfide linkages in the spike protein were reduced with DTT and treated with iodoacetamide to alkylate cysteine residues prior to digestion with PNGaseF and trypsin.

Method Conditions

General conditions for the chromatography and mass spectrometry are listed in Tables 1, 2, and 3. Cyclic IMS methods using five passes of the cyclic device were optimized previously for tri- and disaccharide fragments using O-linked glycopeptides from hemopexin. O-glycopeptides from the SARS-CoV-2 spike protein were quadrupole isolated) prior to collisional activation in the trap region of the Cyclic IMS system (shown in Figure 1). Fragments of the glycopeptides were separated using the cyclic ion mobility cell. Methods were optimized such that trisaccharide fragments passed through the device five times before ejection and detection.
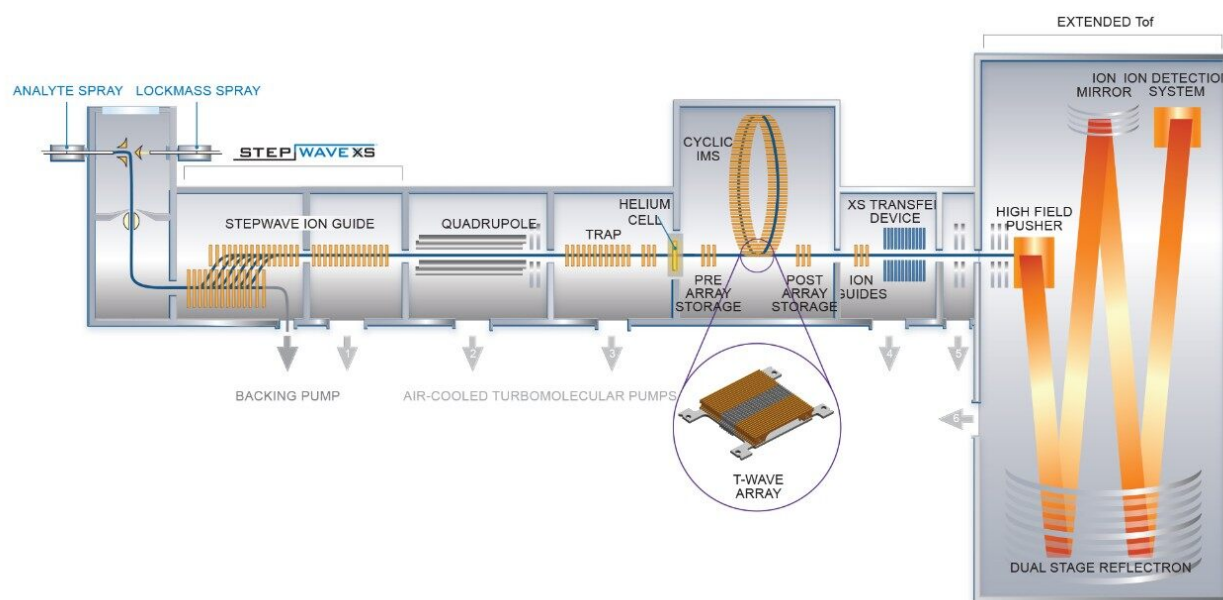

Figure 1. Instrument schematic of the SELECT SERIES Cyclic IMS.

LC Conditions

LC system: ACQUITY UPLC M-Class

Detection: SELECT SERIES Cyclic IMS Mass Spectrometer

Vials:	QuanRecovery
Column(s):	nanoEase M/Z HSS T3 100 Å, 1.8 µm (75 µm x 15 cm) p/n: 186008816
	nanoEase M/Z Symmetry C_{18} 100 Å, 5 µm (180 µm x 20 mm) trap, p/n: 186008821
Column temp.:	60 °C
Sample temp.:	8 °C
Injection volume:	1–5 µL
Flow rate:	500 nL/min
Mobile phase A:	H_2O, 0.1% FA, 1 ppm citric acid
Mobile phase B:	ACN, 0.1% FA, 1 ppm citric acid

Gradient

Time (min)	Flow (µL/min)	%A	%B	Curve
Trapping: 0.0	15	99	1	6
Trapping: 5.0	15	99	1	6
0.0	0.5	99	1	6
3.0	0.5	90	10	6
63.0	0.5	65	35	6
66.0	0.5	15	85	6
69.0	0.5	15	85	6
69.5	0.5	99	1	6
80.0	0.5	99	1	6

MS Conditions

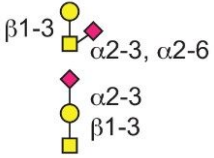|	
MS system:	Cyclic IMS
Ionization mode:	ESI+
Acquisition range:	*m/z* 50–2000
Capillary voltage:	2.8 kV
Collision energy:	30 V–70 V
Cone voltage:	20 V
Number of passes:	5
Racetrack TW height:	22 V
Racetrack TW velocity:	375 m/s
IMS sequence:	Inject, Separate, Eject, and Acquire
IMS cycle time:	100 ms

Data Management

Chromatography software:	MassLynx SCN1016 Release 3
MS software:	Waters Embedded Analyzer Platform for Cyclic IMS Release 4
Informatics:	DriftScope 2.9, modified for Cyclic IMS

Results and Discussion

Targeted ion mobility MS/MS experiments were performed on the observed O-linked glycoforms of tryptic peptide AGC(carbamidomethyl)LIGAEHVDNSYEC(carbamidomethyl)DIPIGAGIC(carbamidomethyl)ASYQTQTNSPR, denoted T56, using fragmentation in the trap region to generate peptidic and oxonium ion fragments prior to the ion mobility cell. The scaleable resolution of the cyclic ion mobility (cIM) instrument was leveraged to separate isomeric glycan structures following release by collisional activation. Two methods were developed previously for 1 and 5 passes of the cyclic device for trisaccharide fragments and were used to separate isomeric forms of the HexNAc-Hex and HexNAc-Hex-NeuAc oxonium ion fragments of T56-HexNAc(1)-Hex(1)-NeuAc(1), T56-HexNAc(1)-Hex(1)-NeuAc(2), T56-HexNAc(2)-Hex(1)-NeuAc(1), T56-HexNAc(2)-Hex(2)-NeuAc(1), and T56-HexNAc(2)-Hex(2)-NeuAc(2). These peptides are listed in Table 1 as glycopeptides A-D.

Precursor name	Glycopeptide ID	Observed drift time of NeuAc-Hex-HexNAc (ms)	Assigned oxonium ion structures
T56-HexNAc(1)-Hex(1)-NeuAc(1)	A	73.8	NeuAc α2-3, Gal β1-3, GalNAc
T56-HexNAc(1)-Hex(1)-NeuAc(2)	B	68.7 70.4 73.8	β1-3 Gal, NeuAc α2-3, α2-6 on HexNAc; NeuAc α2-3, Gal β1-3, GalNAc
T56-HexNAc(2)-Hex(2)-NeuAc(1)	C	73.2 73.8 80.2	NeuAc α2-6, Gal β1-4, GlcNAc; NeuAc α2-3, Gal β1-3, GalNAc; NeuAc α2-3, Gal β1-4, GlcNAc
T56-HexNAc(2)-Hex(2)-NeuAc(2)	D	73.2 73.8 80.2	NeuAc α2-6, Gal β1-4, GlcNAc; NeuAc α2-3, Gal β1-3, GalNAc; NeuAc α2-3, Gal β1-4, GlcNAc

■ N-Acetylgalactosamine (GalNAc)
■ N-Acetylglucosamine (GlcNAc)
■ HexNAc (either GalNAc or GlcNAc)
◯ Galactose (Gal)
◆ N-Acetylneuraminic acid (NeuAc)

Table 1. O-glycopeptides studied by LC-MS-IM-MS and the observed drift times of the observed arrival time distributions of the m/z 657 Neu-Ac-Hex-HexNAc oxonium ion.

Chromatographic peaks yielding a triply charged unmodified T56 peptide as a fragment were integrated in DriftScope 2.9 and the mobility information was extracted. In general, 2–4 chromatographic peaks were observed for each glycoform, indicating that some isomeric heterogeneity can be separated by reverse phase chromatography. Collision induced fragmentation of the T56 glycopeptides generally yielded a mixture of peptidic, oxonium, and peptide + glycan fragments. Figure 2 shows MS/MS spectra for A) glycopeptide A and B) glycopeptide C, demonstrating a high yield of large oxonium ion fragments, particularly for

glycopeptide C. Observation of the HexNAc(2)-Hex(2)-NeuAc(1) oxonium ion fragment at *m/z* 1022 confirms the presence of a single pentasaccharide glycoform rather than separate tri- and disaccharide components localized at two different sites. Oxonium ions at *m/z* 528 and 819 confirm the presence of extended core 1 structures.
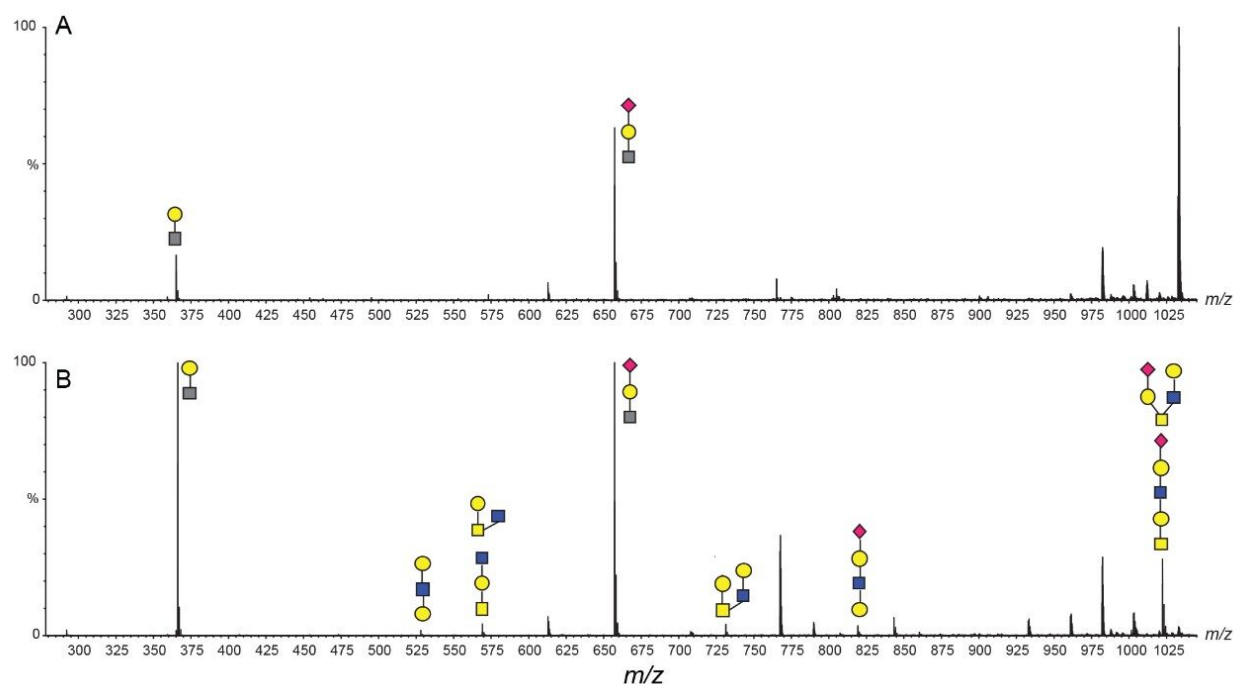

Figure 2. Zoomed MS/MS spectra of A) T56-HexNAc(1)-Hex(1)-NeuAc(2) (glycopeptide A) and B) T56-HexNAc(2)-Hex(2)-NeuAc(1) (glycopeptide C) between m/z 250 and 1050, highlighting several observed oxonium ion fragments that confirm both extended core 1 and core 2 structures are present for glycopeptide C. HexNAc residues are shaded in a gray square if the sugar is ambiguous (GalNAc or GlcNAc); all other colors correspond to SNFG notation (symbol nomenclature for glycans).

Single pass and multi pass (5 cycles) ion mobility experiments were undertaken for the HexNAc-Hex-NeuAc oxonium ion. The resulting arrival time distributions are shown in Figure (3B-D) for glycopeptide C using variable collision energies. Using a single cyclic IM pass, two mobility populations are evident; using five passes, the more compact species further separates into two partially resolved isomers. Collisional cross sections were calculated for the 1 pass using the Major Mix IMS calibration standard as a reference. The two arrival time distributions observed with 1 pass have calculated collisional cross sections of 234.1 and 245.4 Å^2, respectively. Guttman *et. al.* previously measured CCS values for NeuAcα2-6Galβ1-4GlcNAc and NeuAcα2-3Galβ1-4GlcNAc as 236.9 Å^2 and 247.2 Å^2, respectively.[12] The most elongated measured HexNAc-Hex-NeuAc

isomer is in good agreement with a NeuAcα2-3Galβ1-4GlcNAc. Using 5 passes of the cyclic mobility device, the single mobility peak observed at 234.1 Å² was resolved into two substructures. Previous studies have shown alpha 2–6 sialic acid linkages to be more stable than alpha 2–3 linkages; thus, experiments in which the collision energy was varied are shown in Figure 3 (B–D).[17] Higher collision energy yields an increase in the relative abundance of the most compact of the three isomers resolvable with 5 passes. Core 1 type structures (NeuAcα2-3Galβ1-3GalNAc) are the most commonly observed for O-linked glycopeptides; thus, the more elongated and less stable of the compact isomers in the 5-pass separation is consistent with a NeuAcα2-3Galβ1-3GalNAc, type 1 core structure. The more compact and less stable is assigned to a NeuAc α2-6Galβ1-4GlcNAc, which would be expected from fragmentation of a core 2 structure.
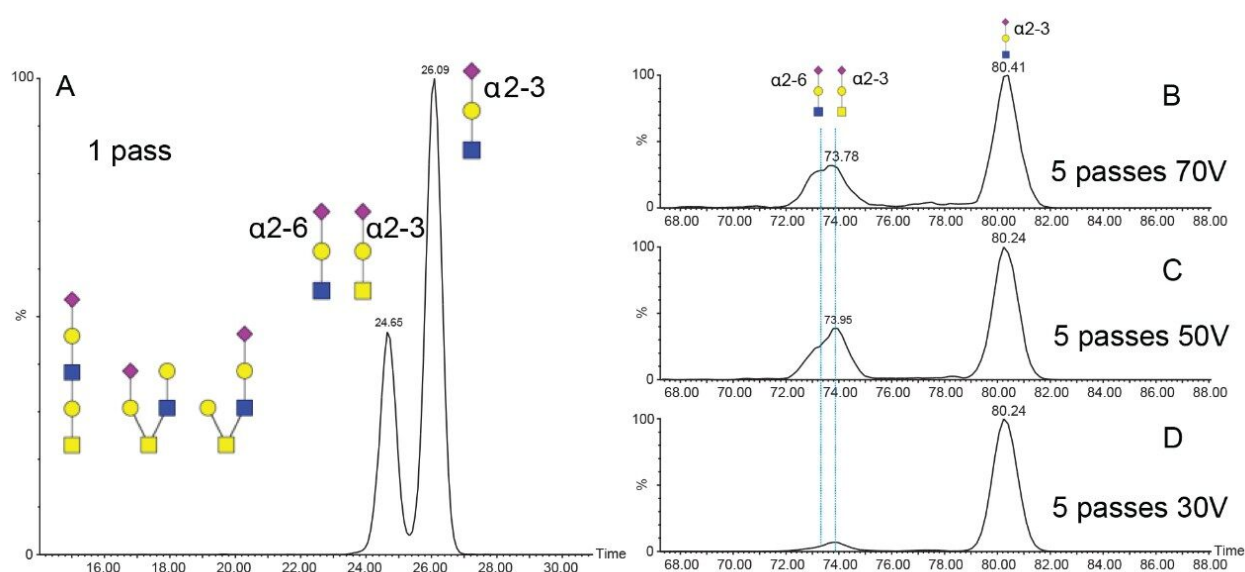

Figure 3. Arrival time distributions for the m/z 657 oxonium ion from glycopeptide C using A) 1 pass of the cyclic device and B–D) 5 passes of the cyclic device with 70, 50, and 30 V activation in the trap to generate the oxonium ion fragments. The additional four passes are necessary for discrimination of the NeuAcα2-6Gal β1-4GlcNAc and NeuAcα2-3Galβ1-3GalNAc isomeric fragments.

Five pass cyclic ion mobility was applied to the *m/z* 657 (HexNAc-Hex-NeuAc) oxonium ion from the remaining glycopeptides and is shown in Figure 4. The glycopeptides having two HexNAc residues (C and D) yielded similar ion mobility arrival time distributions (ATDs), suggesting a mixture of extended core 1 and core 2 structures, as shown in Figure 3. The simplest of the glycopeptides, having a single HexNAc, Hex, and NeuAc residue (A), formed a single mobility distribution at 73.8 ms, which was assigned to NeuAcα2-3Galβ1-3GalNAc for the T56-HexNAc(2)-Hex(2)-NeuAc(1) glycopeptide and is consistent with known glycobiology.

In contrast, separation of the *m/z* 657 oxonium ion from the T56-HexNAc(1)-Hex(1)-NeuAc(2) glycopeptide resulted in two additional mobility peaks at 68.7 and 70.4 ms. These are likely branched structures in which the NeuAc residue is linked to the GalNAc by either 2–3 or 2–6 linkages. These peaks were not observed for glycopeptide D, suggesting sialylation of the core GalNAc residue (Figure 4B) is rare for the T56 glycopeptide and occurs only on the simple core 1 structure.
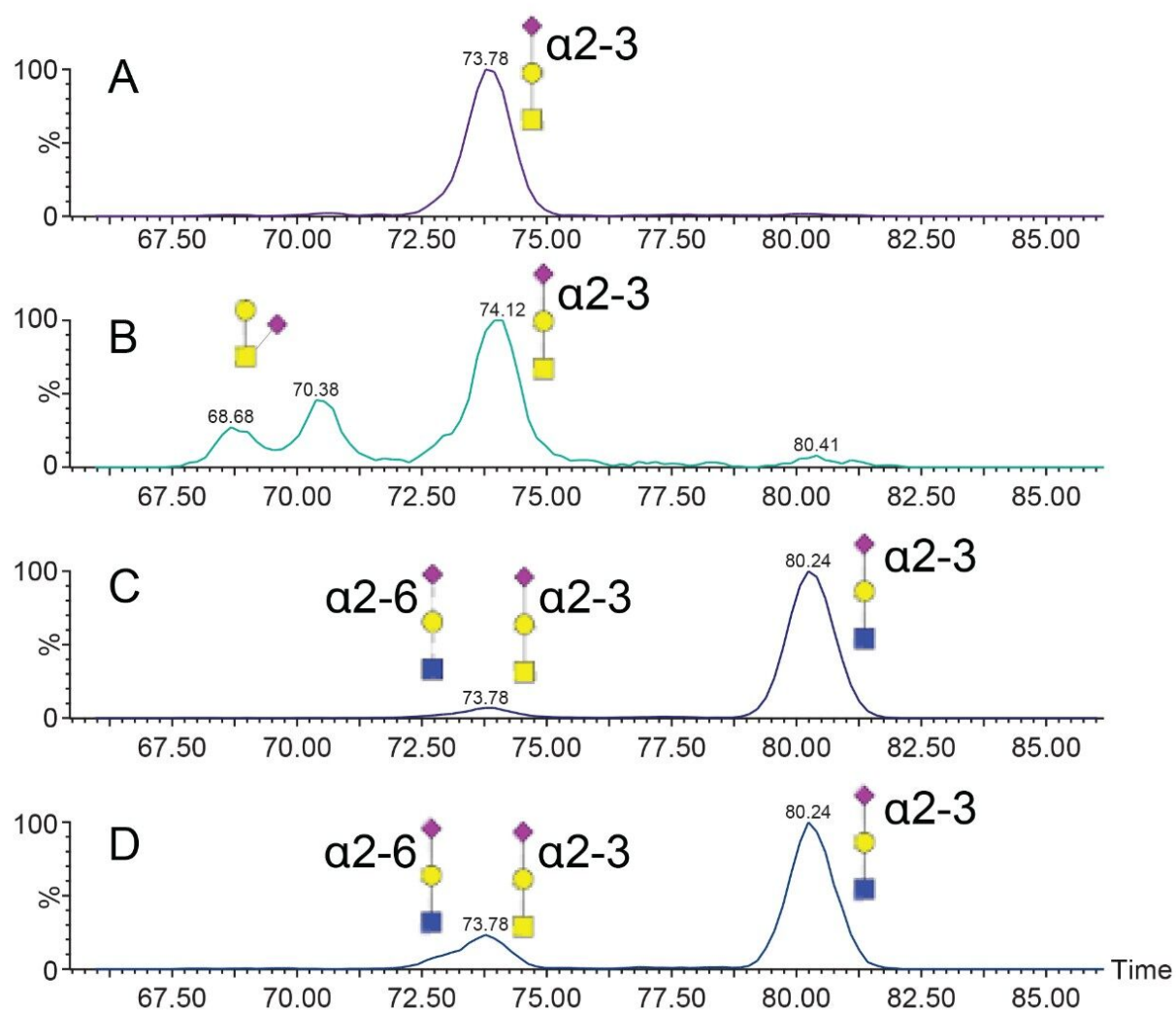

Figure 4. Arrival time distributions of the m/z 657 oxonium ion from A) glycopeptide A, B) glycopeptide B, C) glycopeptide C, and D) glycopeptide D.

Conclusion

The SELECT SERIES Cyclic IMS is demonstrated to show separation of the oxonium ion fragments of SARS-CoV-2 spike protein O-glycopeptides, providing site-specific and linkage specific information about the glycopeptide sugar structures. The high sensitivity, scalable ion mobility resolution, and unique geometry of the Cyclic IMS instrument allows for targeted CID-cIM-MS experiments of low-level O-glycopeptides, affording site-specific structural characterization of highly heterogeneous proteins. This strategy provides improved detail for the ongoing, global effort to characterize, understand, and ultimately target the SARS-CoV-2 virus.

References

1. Johns Hopkins Resource Center. COVID-19 Case Tracker. https://gisanddata.maps.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6.

2. Andersen K.G., Rambaut A., Lipkin W.I. *et al*. The proximal origin of SARS-CoV-2. *Nat Med* 2020; 26:450–452.

3. Tortorici AM, Veesler D. Chapter Four–Structural insights into coronavirus entry, Editor(s): Félix A. Rey, Advances in Virus Research, Academic Press, 2019, 105:93-116, ISSN 0065-3527, ISBN 9780128184561.

4. Walls, AC., Park YJ, Tortorici A, Wall A, McGuire AT, Veesler D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein, *Cell*, 2020;181(2):281-292.e6, ISSN 0092–8674.

5. Clark GF. The Role of Glycans in Immune Evasion: The Human Fetoembryonic Defence System Hypothesis revisited, *Molecular Human Reproduction*. 2014 Mar; 20(3):185–199.

6. Chang D, Zaia J. Why Glycosylation Matters in Building a Better Flu Vaccine. *Molecular & Cellular Proteomics* 2019 Dec; 18 (12):2348–2358.

7. Khatri K, Klein JA, White MR, Grant OC, Leymarie N, Woods RJ, Hartshorn KL, Zaia J. Integrated Omics and Computational Glycobiology Reveal Structural Basis for Influenza A Virus Glycan Microheterogeneity and Host Interactions. *Molecular & Cellular Proteomics* 2016 Jun; 15 (6):1895–1912.

8. Kumar S, Maurya VK, Prasad AK, Bhatt MLB, Saxena SK. Structural, Glycosylation and Antigenic Variation

between 2019 Novel Coronavirus (2019-nCoV) and SARS Coronavirus (SARS-CoV). *Virusdisease*. 2020;31(1):13–21.

9. Watanabe Y, Allen JD, Wrapp D, McLellan JS, Crispin M. Site-Specific Glycan Analysis of the SARS-CoV-2 Spike. *Science*. 2020;eabb9983.

10. Shajahan A, Supekar NT, Gleinich AS, Azadi P. Deducing the N- and O- Glycosylation Profile of the Spike Protein of Novel Coronavirus SARS-CoV-2. *Glycobiology.* 2020; cwaa042. https://doi.org/10.1093/glycob/cwaa042.

11. Sanda M, Morrison L, and Goldman R. N- and O-Glycosylation of the SARS-CoV-2 Spike Protein. *Anal. Chem*. 2021.

12. Guttman M and Lee KK. Site-Specific Mapping of Sialic Acid Linkage Isomers by Ion Mobility Spectrometry. *Anal. Chem*. 2016; 88(10):5212–5217.

13. Jin C, Harvey DJ, Struwe WB, and Karlsson NG. Separation of Isomeric O-Glycans by Ion Mobility and Liquid Chromatography-Mass Spectrometry. *Anal. Chem*. 2019. 91(16): 10604–10613.

14. Harvey DJ, Seabright GE, Vasiljevic S, Crispin M, and Struwe WB. Isomer Information from Ion Mobility Separation of High-Mannos Glycan Fragments. *J. Am. Soc. Mass. Spectrom.* 2018; 29(5): 972–988.

15. Barroso A, Giménez A, Konijnenberg A, Sancho J, Sanz-Nebot V, Sobott F. Evaluation of Ion Mobility for the Separation of Glycoconjugate Isomers Due to Different Types of Sialic Acid Linkage, at the Intact Glycoprotein, Glycopeptide and Glycan Level. *J. Proteom*. 2018. 173:22–31.

16. Hofmann J, Hahm HS, Seeberger PH, and Pagel K. Identification of Carbohydrate Anomers using Ion Mobility-Mass Spectrometry. *Nature*. 2015. 526:241-244.

17. Depland AD, Renois-Predelus G, Schindler B, Compagnon I. Identification of Sialic Acid Linkage Isomers in Glycans using Coupled InfraRed Multiple Photon Dissociation (IRMPD) Spectroscopy and Mass Spectrometry. *Int. J. Mass Spectrom*. 2018; 434:64–69.

Featured Products

SELECT SERIES Cyclic IMS <https://www.waters.com/135021297>

ACQUITY UPLC M-Class System <https://www.waters.com/134776759>

Ion Mobility Mass Spectrometry <https://www.waters.com/134656158>

720007081, Revised January 2021