Distinct Core Glycan and O-Glycoform Utilization of SARS-CoV-2 Omicron Variant Spike Protein RBD Revealed by Top-Down Mass Spectrometry

David S. Roberts,[1]* Morgan Mann,[2] Brad H. Li,[3] Donguk Kim,[3] Allan R. Brasier,[2,4] Song Jin,[1] Ying Ge[1,3]*

[1]Department of Chemistry, University of Wisconsin-Madison, Madison, WI 53706, USA.

[2]Department of Medicine, School of Medicine and Public Health, University of Wisconsin-Madison, Madison, WI 53705, USA.

[3] Department of Cell and Regenerative Biology, University of Wisconsin Madison, WI 53705, USA.

[4]Institute for Clinical and Translational Research, University of Wisconsin-Madison, Madison, WI 53705, USA.

*Correspondence should be addressed to Y.G. (ying.ge@wisc.edu) and D.S.R. (dsroberts@wisc.edu)

Abstract

The SARS-CoV-2 Omicron (B.1.1.529) variant possesses numerous spike (S) mutations particularly in the S receptor-binding domain (S-RBD) that significantly improve transmissibility and evasion of neutralizing antibodies. But exactly how the mutations in the Omicron variant enhance viral escape from immunological protection remains to be understood. The S-RBD remains the principal target for neutralizing antibodies and therapeutics, thus new structural insights into the Omicron S-RBD and characterization of the post-translational glycosylation changes can inform rational design of vaccines and therapeutics. Here we report the molecular variations and O-glycoform changes of the Omicron S-RBD variant as compared to wild-type (WA1/2020) and Delta (B.1.617.2) variants using high-resolution top-down mass spectrometry (MS). A novel O-glycosite (Thr376) unique to the Omicron variant is identified. Moreover, we have directly quantified the Core 1 and Core 2 O-glycan structures and characterized the O-glycoform structural heterogeneity of the three variants. Our findings reveal high resolution detail of Omicron O-glycoforms and their utilization to provide direct molecular evidence of proteoform alterations in the Omicron variant which could shed light on how this variant escapes immunological protection.

Introduction

The SARS-CoV-2 Omicron (B.1.1.529) variant has been classified by the World Health Organization (WHO) as a variant of concern (VOC) due to its significantly increased transmissibility, significant evasion of neutralizing antibodies from convalescents or vaccines, and higher risk of eluding testing.[1, 2] Moreover, recent clinical data showed that this highly mutated Omicron variant causes higher rates of reinfection and rampant breakthrough infections with drastically different clinical outcomes as compared to wild-type (WT, WA1/2020) and Delta (B.1.617.2) variants, despite mRNA vaccine booster dose.[3] Strikingly, the Omicron variant possess an alarming number of mutations (> 30), including 15 site mutations in the S receptor-binding domain (S-RBD) as compared to the WT strain. But how the mutations in the Omicron variant enhance viral escape from immunological protection remains to be understood.

Given that the S-RBD is the principal target for neutralizing antibodies and other therapeutics,[4] and that glycosylation plays critical roles in host receptor ACE2 binding and function,[5-8] it is crucial to decipher the glycoform changes of the Omicron S-RBD as compared to WT and Delta. Although S protein N-glycosylation has been characterized in detail[9, 10] with ongoing efforts to understand the impact of N-glycosylation for vaccine development,[11] characterization of O-glycosylation is challenging[12, 13] due to the large microheterogeneity and structural diversity of O-glycans leading to multiple O-glycoforms.[14, 15] To address these challenges, we have recently developed a hybrid top-down mass spectrometry (MS) approach[16] for comprehensive characterization of O-glycoforms, along with other post-translational modifications (PTMs), to enable proteoform analysis[17, 18] of complex glycoproteins. Because of the rapid mutation and spread of emerging SARS-CoV-2 variants, there is an urgent need for

accurately distinguishing S-RBD variants and elucidating their post-translational glycosylation changes to bridge the knowledge gap between genomic changes and their clinical outcomes.

Here we report the first analysis of the sites and O-glycoform structure differences of the Omicron SARS-CoV-2 S-RBD variant compared to the WT and Delta variant, by top-down MS. Not only has a new O-glycan site been observed, but also top-down O-glycoform quantification revealed significant enhancement of Core 2 type O-glycan structures for Omicron, as compared to the WT or Delta variants.

Results and Discussion

We used HEK293 expressed S-RBD protein arising from WT (WA1/2020), Delta (T478K), and Omicron (BA.1) variants for all the top-down MS analysis. The mutational differences inherent to the Delta and Omicron variants, as compared to the WT strain, are especially pronounced in their RBDs (Fig. 1A). To elucidate the molecular sequence and O-glycans of the various S-RBDs, we removed the N-glycans from the S-RBD using a PNGase F treatment (see Supplementary Information)[16] to minimize the interference posed by N-glycan heterogeneity (Fig. 1B). N-glycans removal yielded a ~10 kDa decrease in the molecular weight by SDS-PAGE as compared to the neat S-RBD. The resulting S-RBD O-glycoforms were resolved by an ultrahigh resolution 12 T Fourier transform ion cyclotron (FTICR)-MS (Fig.1C). Notably, the Omicron variant shows drastic differences in its O-glycosylation profile, as compared to the other variants (Fig. S1-S3). We utilized a timsTOF Pro capable of trapped ion mobility spectrometry (TIMS)-MS[19] to separate and analyze the various S-RBD O-glycan structures for detailed glycoform characterization following the N-glycan removal (Fig. 2). To characterize the glycan structures and occupancy of the highly mutated Omicron variant, we performed specific isolation of the individual S-RBD O-glycoforms. Focusing on the most abundant O-glycoform (26+ charge state, 1069.4.3 *m/z*) as a

specific example, collisionally activated dissociation (CAD) fragment ions were analyzed in targeted protein analysis mode using MASH Explorer[20]. The top-down MS/MS spectra were obtained along with ion mobility separation of the various O-glycan structures to overcome the microheterogeneity inherent to O-glycans analysis (Fig. 2B). Soft TIMS cell activation parameters enabled detailed neutral loss mapping of the isolated S-RBD proteoform and revealed a Core 1 (Galβ1-3GalNAc-Ser/Thr) O-glycan with a GalNAcGal(NeuAc)$_2$ structure (Fig. 2C). Taking advantage of the TIMS separation, we could accurately assign the specific isolated precursor to a single parent O-glycan structure, thereby overcoming mass degeneracy inherent to glycan assignment.[21] Moreover, this TIMS-MS approach revealed multiple S-RBD glycoforms with Core 1 and Core 2 (GlcNAcβ1-6(Galβ1-3)GalNAc-Ser/Thr) O-glycan structures across the three S-RBD variants (Fig. S4). Interestingly, the Omicron variant S-RBD presented different O-glycan patterns as compared to the WT and Delta variants.

Table 1. Summary of the S-RBD variants O-glycoform relative abundances.

O-glycoform structure	Cr.[a]	WT (%)	Delta (%)	Omicron (%)
GalNAcGalNeuAc	1	1.3	1.4	2.3
GalNAcGal(NeuAc)$_2$	1	80.9	86.2	69.0
GalNAc(GalNeuAc)(GlcNAcGal)[b]	2	5.8	3.8	5.4
GalNAc(GalNeuAc)(GlcNAcGalFuc)[b]	2	5.2	3.2	13.2
GalNAc(GalNeuAc)(GlcNAcGalNeuAc)	2	6.8	5.3	10.0

[a] O-glycan core type. [b] Combined abundance including O-glycoforms with a N-terminal acetylation (WT, Delta) or N-terminal trimethylation (Omicron).

We then further characterized the S-RBD WT, Delta, and Omicron O-glycosylation patterns to reveal all the structural O-glycoform alterations between the variants (Fig. 3). Interestingly, we observed major O-glycan microheterogeneity changes in Omicron, as compared to the WT or Delta variants. In particular, we found significantly enhanced Core 2 O-glycan

structure abundances for Omicron with pronounced expression for multiply sialylated GalNAc(GalNeuAc)(GlcNAcGalNeuAc) and fucosylated GalNAc(GalNeuAc)(GlcNAcGalFuc) structures. The striking molecular abundance differences observed in Omicron as compared to the WT or Delta variants are summarized in Table 1. The relative abundance of Core 1 to Core 2 S-RBD O-glycan structures for the Omicron variant was roughly 71:29, with the Core 1 GalNAcGal(NeuAc)$_2$ being the most abundant O-glycoform (~69% relative abundance). Interestingly, the WT and Delta variants show a strong bias toward Core 1 type O-glycan structures, with more than 80% of its O-glycoform abundance corresponding to the Core 1 GalNAcGal(NeuAc)$_2$ structure. The Omicron variant was found to possess a Core 2 type GalNAc(GalNeuAc)(GlcNAcGalFuc) structure that accounts for more than 13% of its total O-glycoform composition (Fig. 3 and S5). Moreover, we characterized the abundant (10 %) Core 2 GalNAc(GalNeuAc)(GlcNAcGalNeuAc) structure for the Omicron variant (Fig. 3 and S6). These particular Core 2 structures were also found in the WT and Delta S-RBD variants but at much lower (5-7%) relative abundances. The high-resolution intact S-RBD glycoform characterization shown in Fig. 3 demonstrates the distinct advantages of this top-down MS approach over glycopeptide-based bottom-up MS approaches.[22, 23]

We then further investigated the glycosites and their microheterogeneity between the S-RBD variants. Fig. 4 shows a specific example of top-down MS/MS O-glycosite determination using the highly abundant Core 1 GalNAcGal(NeuAc)$_2$ structure. Detailed top-down MS/MS analysis of the S-RBD O-glycoforms revealed the presence of a new O-glycosite (Thr376) unique to the Omicron variant (Fig. 4A). Fascinatingly, all detected S-RBD O-glycans for the WT and Delta variants were confidently assigned solely to Thr323 (Fig. 4B and 4C), which agrees with previous studies on WT S O-glycosylation.[24] Representative CAD fragment ions for the WT (b_7^{1+}

and b_{173}^{11+}) and Delta (b_7^{1+} and b_{233}^{12}) confidently localized the O-glycosite to Thr323 (Fig. 4b-c and S7). Given the smaller number of mutations present on Delta as compared to Omicron, it is not surprising that the O-glcyosite Thr323 was conserved between Delta and WT variants. On the other hand, the Omicron variant yielded both the familiar Thr323 O-glycosite (b_5^{1+} and b_{18}^{2+}) and a new Thr376 O-glycosite (b_{60}^{12+} and b_{52}^{5+}) corresponding to the GalNAcGal(NeuAc)$_2$ O-glycoform that is simultaneously occupied (Fig. 4D). This Thr376 O-glycosite is conveniently n + 3 adjacent to a proline at residue 373, which is consistent with previous reports of increased O-glycosylation frequency near proline.[25] This particular Pro373 is a site-specific mutation unique to the Omicron variant and likely is the reason for this new O-glycosite. We note that the site occupancy of the Thr376 site is low(< 30%) relative to the Thr323 and was only confidently assigned for the abundant GalNAcGal(NeuAc)$_2$ O-glycoform, although we suspect other Core 2 O-glycoforms may also possess the Thr376 modification. It should be noted that recombinant fragments of S protein may show differential glycosylation when the S-RBD is expressed as a monomer, therefore care is needed when assigning O-glycans between different protein sources.[26] Moreover, although the S-RBD O-glycoforms assigned for the variants are specific to HEK293 derived S-RBD, the HEK293 expression model has been shown to reflect glycosylation sites expected for the viron.[27-29]

Conclusion

In summary, we report the first analysis of the O-glycoform structural heterogeneity of the S-RBD found in the SARS-CoV-2 Omicron variant. We observed significant enhancement in the utilization of Core 2 type O-glycoforms for the Omicron variant as compared to WT or Delta. Moreover, we identified and characterized a novel Thr376 O-glycosite unique to Omicron S-RBD. This top-down MS approach is complimentary to traditional structural methods such as X-ray

crystallography and cryoEM, which are not amenable for direct glycan structural analysis due to the inherent flexibility and heterogeneity of oligosaccharides,[30, 31] and provides unmatched resolution for the characterization of SARS-CoV-2 S-RBD proteoforms. Importantly, this top-down MS approach can be leveraged to resolve the protein level changes of S mutations in emerging SARS-CoV-2 variants to understand changes in their structure-function. Our findings bridge the knowledge gap between S variant genomic alterations and final clinical outcomes with detailed proteoform information, which could shed new light on how Omicron escapes immunological protection.

Acknowledgement

This research is supported by NIH R01 GM117058 (to S.J. and Y.G.). D.S.R. acknowledges the support from the American Heart Association Predoctoral Fellowship Grant No. 832615/David S. Roberts/2021. Y.G. would like to acknowledge NIH R01 GM125085, R01 HL096971, and S10 OD018475. We also thank Michael Greig, Guillaume Tremintin, Conor Mullins, Gary Kruppa, Paul Speir, and Rohan Thakur of Bruker Daltonics for their helpful discussion and provision of the Bruker timsTOF Pro used in this work.

Notes

The authors declare no competing financial interest.

Figures

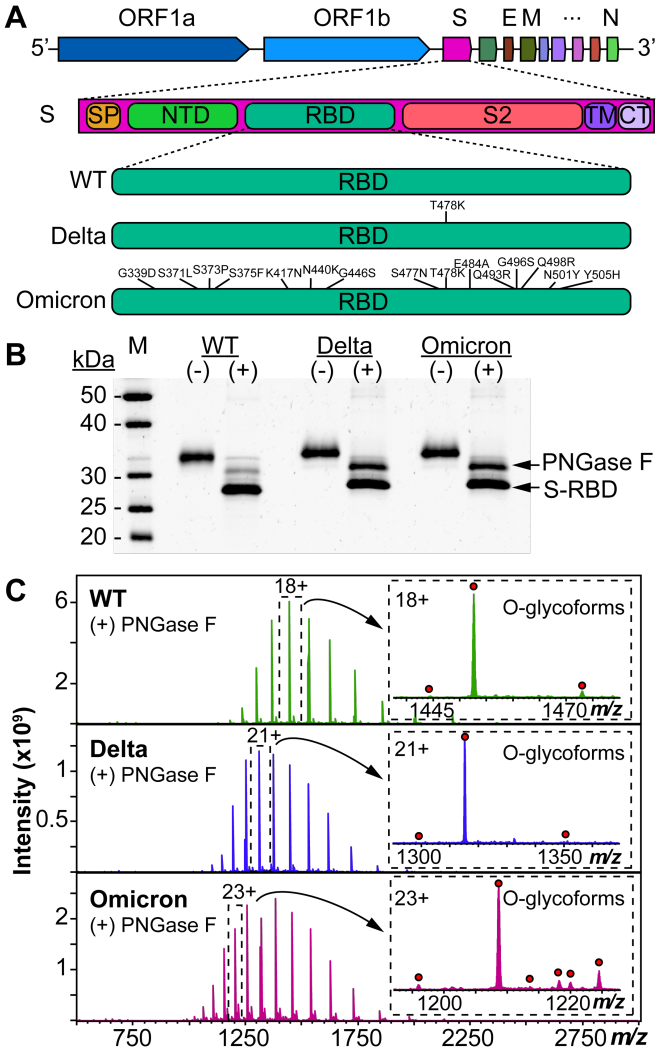

Figure 1. Protein mutational mapping and high-resolution top-down MS of S-RBD arising from the WT, Delta, and Omicron variants. (A) Architecture of SARS-CoV-2 genome and illustration of the protein sequence changes for the S-RBD variants. (B) SDS-PAGE of the S-RBD variants before (−) and after (+) PNGase F treatment. Gel lanes were equally loaded (300 ng), and gel staining was visualized by SYPRO Ruby. (C) Raw MS1 of intact S-RBD proteoforms collected on the 12 T FTICR-MS after PNGase F treatment for the WT, Delta, and Omicron variants. All identified O-glycoforms are annotated in the inset with a red circle.

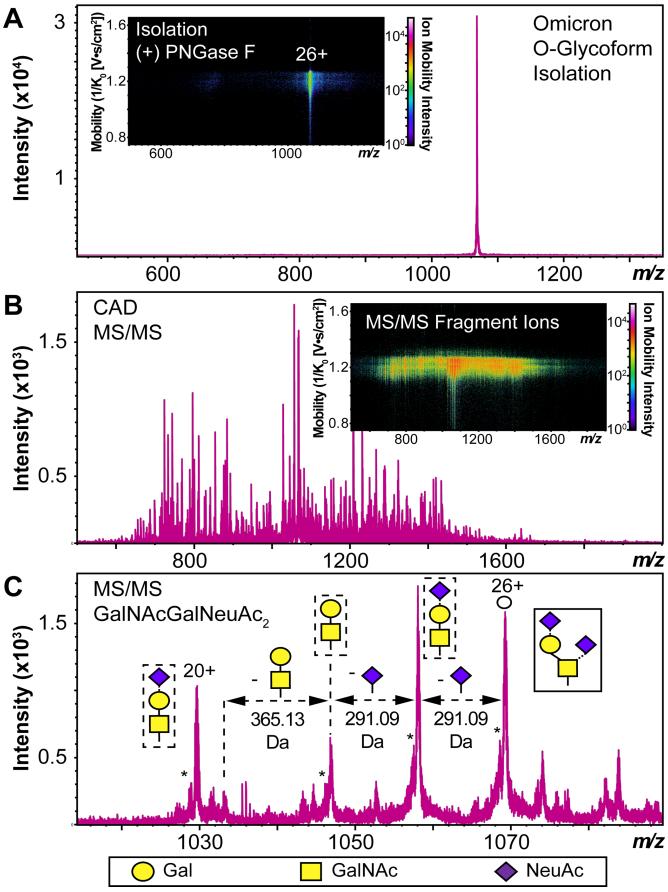

Figure 2. S-RBD O-glycoform analysis by TIMS-MS/MS. (A) Illustration of the isolation of a specific Omicron S-RBD glycoform (z = 26+, centered at 1069.4 m/z) after PNGase F treatment using TIMS-MS. (Inset) Associated ion mobility heat map after isolation of the precursor. (B) Top-down MS/MS of Omicron S-RBD O-glycoform following CAD fragmentation of the isolated proteoform. A collision cell voltage of 22 eV was applied for this specific example. (C) Neutral loss O-glycan mapping of the Omicron S-RBD proteoform. Glycoform characterization reveals the specific S-RBD proteoform to have a Core 1 type GalNAcGalNeuAc2 glycan structure. Assignments of the glycan structures are marked in the spectrum with the legend shown on the bottom. The hollow circle represents the 26+ charge state precursor ion corresponding to the 1069.4 m/z isolation. An asterisk "*" denotes an oxonium ion loss.

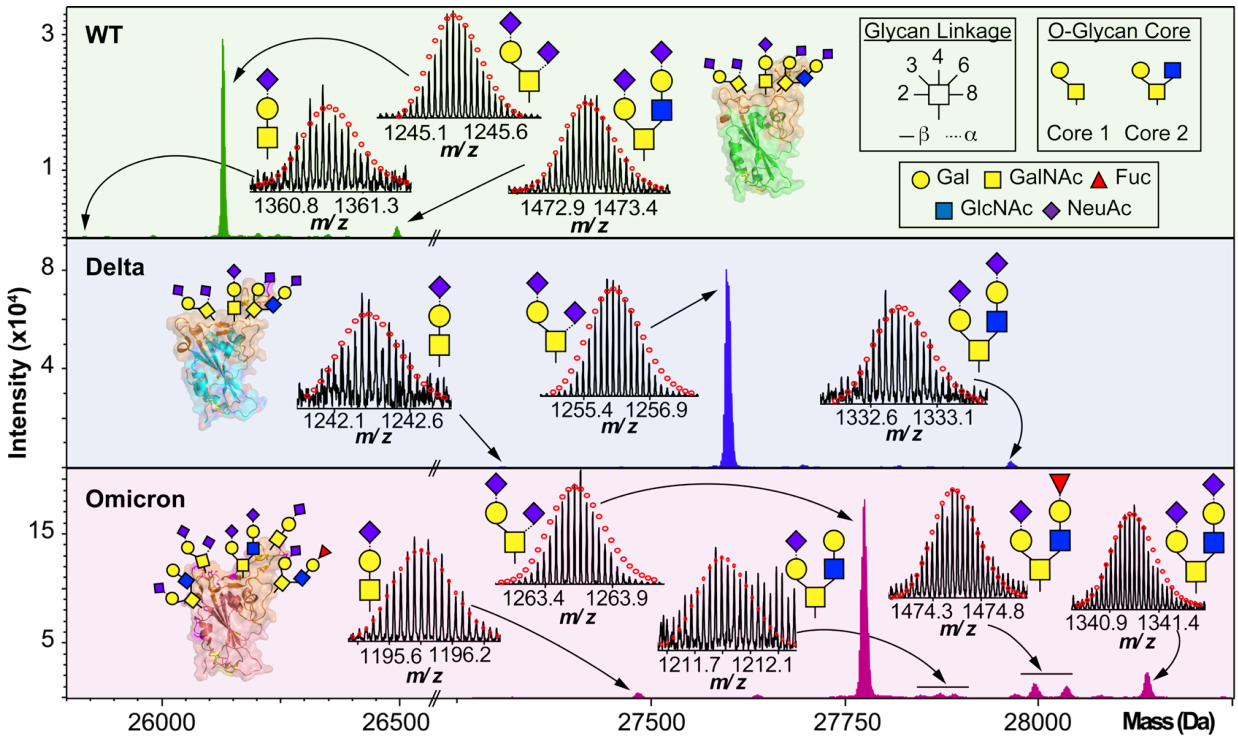

Figure 3. O-glycoform characterization of the S-RBD variants. Deconvoluted MS of the S-RBD proteoforms are shown with all major O-glycan assignments for the WT (green), Delta (blue), and Omicron (pink) variants. Insets show the isotopically and baseline resolved S-RBD proteoform MS spectra allowing for accurate determination of individual intact glycoforms. MS data were collected on the 12 T FTICR-MS. All individual ion assignments are within 1 ppm from the theoretical mass and theoretical isotopic distributions are indicated by the red dots. Assignments of the glycan structures are marked in the spectrum with the legend shown as an inset. PDB: 6M0J, 7WBQ, 7WBP.

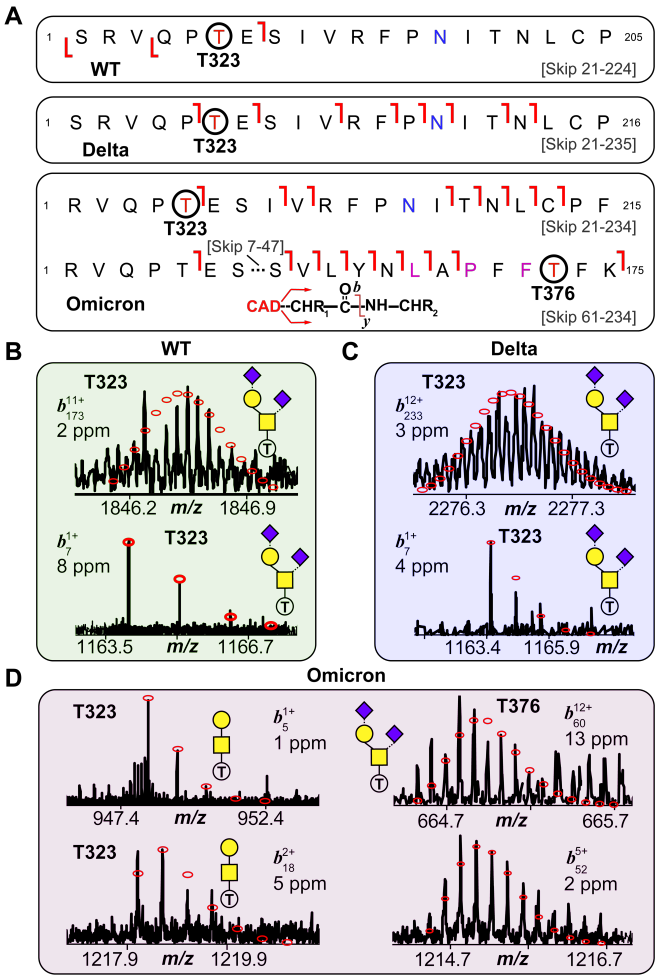

Figure 4. S-RBD O-glycosite localization by top-down TIMS-MS/MS. (A) Fragmentation mapping of the Core 1 type GalNAcGal(NeuAc)2 glycan corresponding to the WT, Delta, and Omicron S-RBD variants. Amino acid sequence (Arg319-Phe541) was based on the entry name P0DTC2 (SPIKE_SARS2) obtained from the UniProtKB sequence database with appropriate amino acid substitutions made for the mutated Delta and Omicron variants. Cleavage of signal peptide tPA following cell expression results in the N-terminal Ser for the WT and Delta variants. The blue N (Asn) denotes deamidation following PNGase F treatment. Specific Omicron residue mutations are denoted with a pink residue. (B-D) Representative top-down MS/MS CAD fragment ions of intact WT (b71+ and b17311+) (B), Delta (b71+ and b22312+) (C), and Omicron (b51+, b182+, b71+, b17311+) (D) variants corresponding to the GalNAcGal(NeuAc)2 O-glycoform shown in (A). WT and Delta variants showed complete O-glycosite occupation at Thr323. The Omicron variant was found to possess both Thr323 (left side; b51+and b182+) and Thr376 (right side; b71+, and b17311+) as O-glycosites corresponding to the same Core 1 O-glycan structure. Theoretical isotopic distributions are indicated by the red dots and mass accuracy errors are listed for each fragment ion.

References

1. VanBlargan, L.A. et al. An infectious SARS-CoV-2 B.1.1.529 Omicron virus escapes neutralization by therapeutic monoclonal antibodies. *Nature Medicine* (2022).
2. Morens, D.M., Taubenberger, J.K. & Fauci, A.S. Universal Coronavirus Vaccines — An Urgent Need. *New England Journal of Medicine* **386**, 297-299 (2021).
3. Kuhlmann, C. et al. Breakthrough infections with SARS-CoV-2 omicron despite mRNA vaccine booster dose. *The Lancet* (2022).
4. Liu, L. et al. Striking Antibody Evasion Manifested by the Omicron Variant of SARS-CoV-2. *Nature* (2021).
5. Harbison, A.M. et al. Fine-tuning the spike: role of the nature and topology of the glycan shield in the structure and dynamics of the SARS-CoV-2 S. *Chemical Science* **13**, 386-395 (2022).
6. Casalino, L. et al. Beyond Shielding: The Roles of Glycans in the SARS-CoV-2 Spike Protein. *ACS Central Science* **6**, 1722-1734 (2020).
7. Zhang, L. et al. Furin cleavage of the SARS-CoV-2 spike is modulated by O-glycosylation. *Proceedings of the National Academy of Sciences* **118**, e2109905118 (2021).
8. Harvey, W.T. et al. SARS-CoV-2 variants, spike mutations and immune escape. *Nature Reviews Microbiology* **19**, 409-424 (2021).
9. Lenza, M.P. 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. *Angewandte Chemie International Edition* **59**, 23763-23771 (2020).
10. Watanabe, Y., Allen, J.D., Wrapp, D., McLellan, J.S. & Crispin, M. Site-specific glycan analysis of the SARS-CoV-2 spike. *Science*, eabb9983 (2020).
11. Huang, H.-Y. et al. Impact of glycosylation on a broad-spectrum vaccine against SARS-CoV-2. *bioRxiv*, 2021.2005.2025.445523 (2021).
12. Rangel-Angarita, V. & Malaker, S.A. Mucinomics as the Next Frontier of Mass Spectrometry. *ACS Chemical Biology* **16**, 1866-1883 (2021).
13. Escobar, E.E. et al. Precision Mapping of O-Linked N-Acetylglucosamine Sites in Proteins Using Ultraviolet Photodissociation Mass Spectrometry. *Journal of the American Chemical Society* **142**, 11569-11577 (2020).
14. Jensen, P.H., Kolarich, D. & Packer, N.H. Mucin-type O-glycosylation – putting the pieces together. *The FEBS Journal* **277**, 81-94 (2010).
15. Riley, N.M., Bertozzi, C.R. & Pitteri, S.J. A Pragmatic Guide to Enrichment Strategies for Mass Spectrometry–Based Glycoproteomics. *Molecular & Cellular Proteomics* **20**, 100029 (2021).
16. Roberts, D.S. et al. Structural O-Glycoform Heterogeneity of the SARS-CoV-2 Spike Protein Receptor-Binding Domain Revealed by Top-Down Mass Spectrometry. *Journal of the American Chemical Society* **143**, 12014-12024 (2021).
17. Tiambeng, T.N. et al. Nanoproteomics enables proteoform-resolved analysis of low-abundance proteins in human serum. *Nature Communications* **11**, 3903 (2020).
18. Brown, K.A., Melby, J.A., Roberts, D.S. & Ge, Y. Top-down proteomics: challenges, innovations, and applications in basic and clinical research. *Expert Review of Proteomics* **17**, 719-733 (2020).
19. Ridgeway, M.E., Bleiholder, C., Mann, M. & Park, M.A. Trends in trapped ion mobility – Mass spectrometry instrumentation. *TrAC Trends in Analytical Chemistry* **116**, 324-331 (2019).
20. Wu, Z. et al. MASH Explorer: A Universal Software Environment for Top-Down Proteomics. *Journal of Proteome Research* **19**, 3867-3876 (2020).
21. Liu, F.C., Cropley, T.C., Ridgeway, M.E., Park, M.A. & Bleiholder, C. Structural Analysis of the Glycoprotein Complex Avidin by Tandem-Trapped Ion Mobility Spectrometry–Mass Spectrometry (Tandem-TIMS/MS). *Analytical Chemistry* **92**, 4459-4467 (2020).

22. Yen, H.-Y. et al. Correlating Glycoforms of DC-SIGN with Stability Using a Combination of Enzymatic Digestion and Ion Mobility Mass Spectrometry. *Angewandte Chemie International Edition* **59**, 15560-15564 (2020).
23. Melby, J.A. et al. Novel Strategies to Address the Challenges in Top-Down Proteomics. *Journal of the American Society for Mass Spectrometry* **32**, 1278-1294 (2021).
24. Shajahan, A., Supekar, N.T., Gleinich, A.S. & Azadi, P. Deducing the N- and O-glycosylation profile of the spike protein of novel coronavirus SARS-CoV-2. *Glycobiology* (2020).
25. Hema Thanka Christlet, T. & Veluraja, K. Database Analysis of O-Glycosylation Sites in Proteins. *Biophysical Journal* **80**, 952-960 (2001).
26. Eldrid, C.F.S., Allen, J.D., Newby, M.L. & Crispin, M. Suppression of O-Linked Glycosylation of the SARS-CoV-2 Spike by Quaternary Structural Restraints. *Analytical Chemistry* **93**, 14392-14400 (2021).
27. Bagdonaite, I. & Wandall, H.H. Global aspects of viral glycosylation. *Glycobiology* **28**, 443-467 (2018).
28. Brun, J. et al. Assessing Antigen Structural Integrity through Glycosylation Analysis of the SARS-CoV-2 Viral Spike. *ACS Central Science* **7**, 586-593 (2021).
29. Sanda, M., Morrison, L. & Goldman, R. N- and O-Glycosylation of the SARS-CoV-2 Spike Protein. *Analytical Chemistry* **93**, 2003-2009 (2021).
30. Li, H., Nguyen, H.H., Ogorzalek Loo, R.R., Campuzano, I.D.G. & Loo, J.A. An integrated native mass spectrometry and top-down proteomics method that connects sequence to structure and function of macromolecular complexes. *Nature Chemistry* **10**, 139 (2018).
31. Chang, V.T. et al. Glycoprotein Structural Genomics: Solving the Glycosylation Problem. *Structure* **15**, 267-273 (2007).