# Structural O-Glycoform Heterogeneity of the SARS-CoV-2 Spike Protein Receptor-Binding Domain Revealed by Native Top-Down Mass Spectrometry

David S. Roberts[1*], Morgan W. Mann[2], Jake A. Melby[1], Eli J. Larson[1], Yanlong Zhu[3,4], Allan R. Brasier[2,5], Song Jin[1], Ying Ge[1,2,3]*

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

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

[3]Human Proteomics Program, School of Medicine and Public Health, University of Wisconsin-Madison, Madison, Wisconsin 53705, USA.

[4]Department of Cell and Regenerative Biology, University of Wisconsin-Madison, Madison, Wisconsin 53705, USA.

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

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

**Abstract**

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) utilizes an extensively glycosylated surface spike (S) protein to mediate host cell entry and the S protein glycosylation is strongly implicated in altering viral binding/function and infectivity. However, the structures and relative abundance of the new O-glycans found on the S protein regional-binding domain (S-RBD) remain cryptic because of the challenges in intact glycoform analysis. Here, we report the complete structural characterization of intact O-glycan proteoforms using native top-down mass spectrometry (MS). By combining trapped ion mobility spectrometry (TIMS), which can separate the protein conformers of S-RBD and analyze their gas phase structural variants, with ultrahigh-resolution Fourier transform ion cyclotron resonance (FTICR) MS analysis, the O-glycoforms of the S-RBD are comprehensively characterized, so that seven O-glycoforms and their relative molecular abundance are structurally elucidated for the first time. These findings demonstrate that native top-down MS can provide a high-resolution proteoform-resolved mapping of diverse O-glycoforms of the S glycoprotein, which lays a strong molecular foundation to uncover the functional roles of their O-glycans. This proteoform-resolved approach can be applied to reveal the structural O-glycoform heterogeneity of emergent SARS-CoV-2 S-RBD variants, as well as other O-glycoproteins in general.

**Introduction**

The novel zoonotic 2019 coronavirus disease (COVID-19) global pandemic[1] has led to more than 102 million reported cases and over 2 million deaths as of February 2021.[2] The causative pathogen of COVID-19, severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), utilizes an extensively glycosylated spike (S) protein that protrudes from the viral surface to bind receptor angiotensin-converting enzyme 2 (ACE2) for cell entry.[3,4] The evolution of the surface S protein is particularly important as it initiates pathogenesis,[5] is the main target of vaccination[6,7] and antibody therapeutic design,[8,9] and a substantial number of predicted mutations in the S protein regional-binding domain (RBD) can potentially enhance ACE2 binding.[10] Moreover, the more than 636,000 viral genomic sequences generated from global molecular epidemiology studies (as of February 28, 2021, www.gisaid.org)[11] reveal the emergence of a number of variants in the S protein, indicating that SARS-CoV-2 may be evolving to acquire selective replication advantages, drug insensitivity and immunological resistance, thereby underscoring the need for new technologies capable of rapid and deep structural profiling of the virus.[12-14]

Importantly, glycans flank the polybasic furin cleavage site of S protein necessary for S stability during biogenesis and influence conformational dynamics by masking protein regions from cleavage during ACE2 binding.[3,15,16] How these different sites are glycosylated and how they influence ACE2 binding are likely to affect cell infectivity and could shield certain epitopes from antibody neutralization.[17-20] The SARS-CoV-2 S protein carries 22 N-glycosylation sequons per protomer which have been characterized in detail.[21,22] However, characterization of O-glycosylation has so far been limited, with only putative O-glycosites reported,[16,22,23] and the glycan microhetereogeneity, including molecular compositions, structures, and relative

3

abundance, remains cryptic due to the challenges in O-glycan analysis.[5,17,24-29] Importantly, individual glycosites can give rise to many glycan structural variants, *i.e.* glycoforms, and these differences in glycan structures, even at a single glycosite, can have critical implications on biological functions.[20,30-32] It is crucial to decipher the intact glycoforms of various O-glycosites recently found in the S-RBD because they are expected to play a role in viral function and viral binding with ACE2.[17,22,23,33] Therefore, new methods capable of the comprehensive structural elucidation of O-glycans are essential for understanding the functional roles of O-glycans on SARS-CoV-2 pathology and provide new avenues for rational therapeutic development.[28,34-36]

Conventional methods to analyze protein glycosylation rely on the bottom-up MS approach of analyzing glycopeptides obtained through enzymatic digestion.[31,37,38] Although the bottom-up MS approach is a high throughput method for global proteome identification of glycosites and N-glycan heterogeneity from complex samples, the relative abundance of various intact glycoforms and the overall post-translational modification (PTM) compositions of different co-appearing 'proteoforms'[39] are lost.[40-42] In contrast, by combining glycoproteomics with the top-down MS approach, which preserves the intact glycoprotein enabling high-resolution proteoform-resolved analysis,[43-45] we could achieve the simultaneous characterization of the molecular structures, the site specificity, and the relative abundance of various glycoforms. Furthermore, by integrating native MS, which has recently emerged as a powerful structural biology tool to study protein structure-function relationships,[46-52] with trapped ion mobility spectrometry (TIMS),[53-55] we can investigate the gas phase structural variants to achieve the direct quantification of individual glycoproteoforms.

Here we developed a method for the molecularly detailed analysis of intact O-glycan proteoforms of the S-RBD by top-down MS (**Figure 1**). For the first time, we harness the capabilities of a hybrid TIMS quadrupole time-of-flight (QTOF) mass spectrometer (timsTOF Pro) (Figure 1C), which provides high resolving power and sensitivity for both selective and comprehensive ion mobility separations of various protein structural variants,[56-59] and the ultrahigh-resolution of a 12T Fourier Transform Ion Cyclotron Resonance (FTICR) MS (Figure 1D), to comprehensively characterize the O-glycoforms of the S-RBD, including the exact glycan structures and relative molecular abundance (Figure 1E). We demonstrate that this native top-down MS approach can provide a high-resolution and wholistic proteoform-resolved landscape of diverse O-glycoforms to enable future structure-function studies of the S-RBD.
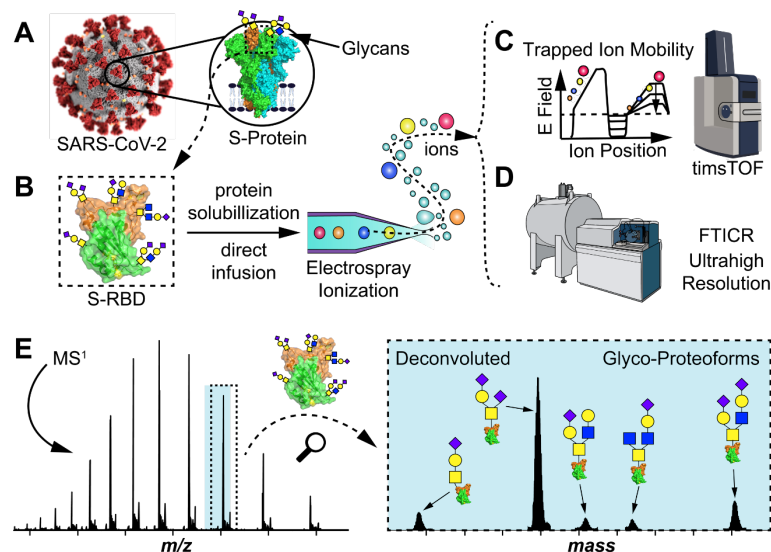


**Figure 1.** High-resolution glycoform characterization of intact S-RBD by top-down MS. (A-D) Illustration of the top-down glycoproteomics workflow for the comprehensive analysis of the S-RBD glycoforms. (A) The SARS-CoV-2 coronavirus features a surface S protein that possesses a glycosylated RBD (highlighted in the dashed box in B). (B) Intact glycoprotein analysis proceeds by directly infusing solubilized S-RBD and electrospraying S-RBD protein ions into either (C) a hybrid trapped ion mobility spectrometry (TIMS) MS device (timsTOF) for ion mobility MS analysis, or (D) an ultrahigh-resolution Fourier transform ion cyclotron (FTICR) MS. (E) Specific isolation of S-RBD proteoforms by top-down MS analysis, illustrating raw MS[1] and corresponding deconvoluted protein spectrum, for structural characterizations of glycoforms. PDB: 6M0J.

5

**Results and Discussion**

**Native TIMS-MS analysis of the S-Protein RBD.** We first prepared a native S-RBD protein expressed from HEK 293 cells to perform native top-down MS for comprehensive glycoform analysis (**Figure 1**). After ensuring reproducible native S-RBD protein sample preparation (**Figure S1**), we used the TIMS, the front-end of the timsTOF Pro (see Methods for details), to achieve high resolution ion mobility separation of protein conformers. TIMS, similar to other quantitative ion mobility techniques, reports the rotationally average collision cross section (CCS) of proteins, which relates to their overall size and shape, and consequently can be used to evaluate changes in three-dimensional structure.[57] However, ion mobility mapping (**Figures 2A**) and MS[1] analysis (**Figure 2B**) reveal enormous complexity in the native S-RBD sample. The diffuse spread of the ion mobility spectrum suggests multiple co-existing S-RBD proteoforms potentially arising from multiple glycans (**Figure 2A**). Closer inspection of the ion mobility and native MS spectra reveals significant glycan heterogeneity likely due to the multiple N- and O-glycans on the native S-RBD (**Figures 2C,D**). The native S-RBD could not be resolved due to overlapping ion signals from all diverse glycoforms/proteoforms carrying different charges, similarly as observed in complex glycoproteins previously.[40,53]
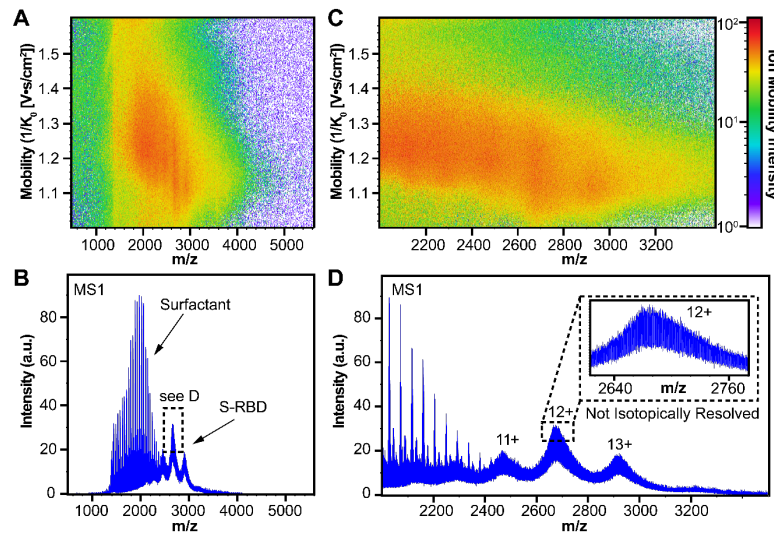
**Figure 2.** Native top-down MS of intact S-RBD expressed from HEK 293 cell line using TIMS-MS. (A-B) Ion mobility heat map (A) and native MS[1] analysis (B) of the S-RBD. (C) Magnified ion mobility heat map for the S-RBD region from (A). The diffuse ion mobility spectrum suggests a diverse spread of multiple protein ions with various collisional cross sections crowded in the native spectrum. (D) Magnified MS[1] from (B) with charge states of the native S-RBD annotated. The inset reveals that the large cluster of various proteoforms of the S-RBD inhibits its isotopic resolution.

The S protein is known to carry many complex-type N-glycans, with two glycosites (Asp331 and Asp343) previously reported and extensively characterized on the S-RBD.[21,22] By contrast, the O-glycans of the S-RBD are less studied and their glycoforms poorly understood, despite the potential insights they may provide to understand viral structure-function relationship.[17,21-23] Since our focus is on O-glycans of the S-RBD, we completely removed the N-glycans from the S-RBD using a native PNGase F treatment (see Methods for details) to minimize the interference posed by the enormous N-glycan heterogeneity (**Figure 3A**). The removal of N-glycans constitutes >10 kDa of molecular weight loss compared to the fully glycosylated S-RBD, as expected from the multiple proteoforms detected by TIMS analysis (**Figures 2A,C**). Fortunately, we found that after complete removal of N-glycans, the remaining O-glycans on the S-RBD were now fully resolvable by native MS (**Figure 3B**). Although the remaining PNGase F was found coexisting with the S-RBD, further optimization of the deglycosylation treatment

7

yielded highly resolved charge states of S-RBD (**Figure 3**C). Deconvoluted MS analysis of the isotopically resolved charge states revealed multiple O-glycoforms with molecular weights ranging from 25.5 kDa to 26.8 kDa (**Figure 3C** inset).
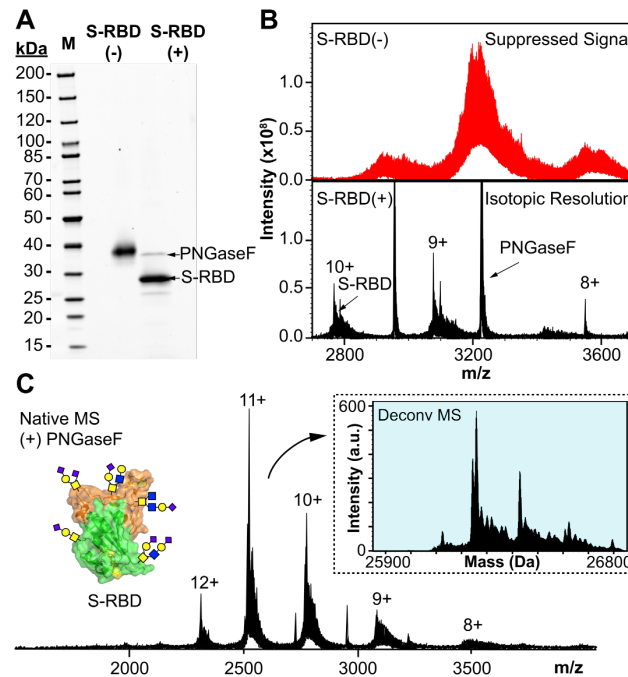


**Figure 3.** Isotopically resolved native top-down MS of intact S-RBD after N-glycan removal. (A) SDS-PAGE of S-RBD before (-) and after (+) PNGase F treatment. The gel lanes were equally loaded (200 ng) and staining was visualized by SYPRO Ruby. (B) Native MS of the S-RBD before (-) and after (+) PNGase F treatment showing improved ion signal and resolution, corresponding to the data shown in (A). (C) Isotopically resolved S-RBD obtained after optimization of PNGase F treatment. The protein charge states are labeled, and the inset shows the corresponding deconvoluted mass spectrum.

We next used TIMS to separate and analyze the various S-RBD glycoforms after the N-glycan removal by taking full advantage of the TIMS front-end of the timsTOF Pro instrument. After careful TIMS tuning and optimization of the ion mobility parameters (**Figure S2**), the timsTOF was sufficiently sensitive to allow high resolution ion mobility separation of intact S-RBD O-glycoforms (**Figure 4A**, **Figure S3**), in contrast to the poorly resolved native S-RBD (**Figure 2**). The TIMS analysis revealed two distinct protein gas-phase conformers separated in regional mobility between 1.1 to 1.45 $1/K_0$. By interrogating each of the resolved mobility regions,

MS[1] analysis revealed that the two gas-phase conformers show shifts in the relative abundance of the S-RBD charge states (**Figure 4B**). The shift in protein charging along with the regional mobility shift together imply possible structural changes which can influence the protein ionization between the mobility regions. This difference in protein conformers was further investigated by plotting the corresponding ion mobility spectrum of the various regions against their calculated CCS values (**Figure 4C**).
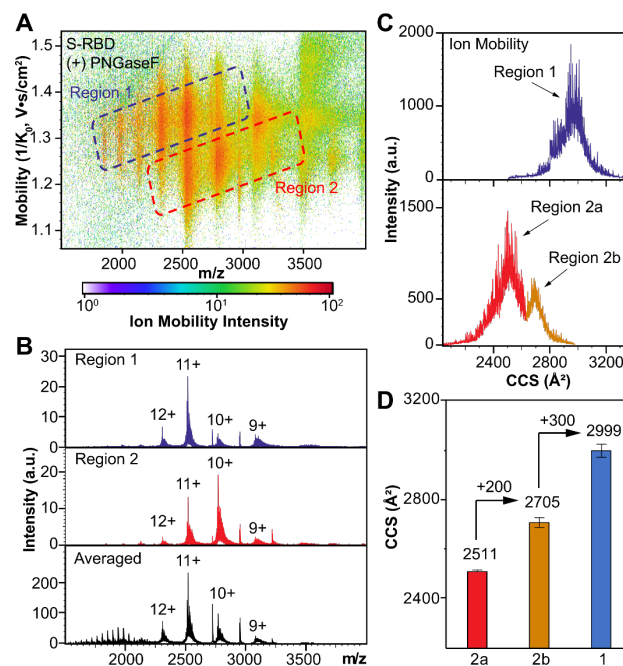


**Figure 4.** Ion mobility separation of intact S-RBD O-glycoforms by native TIMS-MS. (A) Ion mobility heat map of S-RBD after PNGase F treatment reveals distinct gas-phase conformers labeled as Region 1 (blue) and Region 2 (red). (B) Native MS[1] analysis of the various ion mobility regions corresponding to (A). The MS[1] spectrum averaged across the two mobility regions is shown for reference. (C) Ion mobility spectrum corresponding to the data shown in (A). The two distinct peaks in Region 2 are illustrated as Region 2a (red) and Region 2b (orange). (D) Collision cross section (CCS) of the most abundant S-RBD ion (11+ charge state) ion mobility regions corresponding to (C). CCS data (D) are representative of $n = 3$ independent experiments with error bars indicating the standard deviation of the mean.

Applying this method, we further found two distinct S-RBD O-glycoform mobility regions in the mobility Region 2 (**Figure 4C** lower panel). The smallest mobility region of the S-RBD O-glycoforms (Region 2a, $2511 \pm 5$ Å$^2$) is slightly higher than the theoretical value (~2200 Å$^2$)

9

calculated using the IMPACT method[60] from the S-RBD X-ray crystal structure without glycosylation.[61] This discrepancy is likely a result of the influence of the additional O-glycans.[62] CCS measurements of the various ion mobility regions of the S-RBD revealed progressively larger structures between the highlighted mobility regions and further illustrates the separation of the distinct gas-phase conformers (**Figure 4D**). Region 1 (2999±26 Å$^2$) shows the largest calculated CCS value compared to Region 2b (2705±25 Å$^2$) and Region 2a (2511±5 Å$^2$), implying that the S-RBD glycoforms in Region 1 experience additional protein unfolding which increases the overall protein conformer size and signal abundance under electrospray ionization.[63] The variations in S-RBD O-glycoforms CCSs are mobility region specific and consistent with protein conformer changes, illustrating the potential of this approach for investigating the gas-phase structural variations of glycoproteins.

**Comprehensive characterization of S-RBD O-glycoforms by high resolution top-down MS.** Although S-RBD glycoforms can be inferred from the timsTOF analysis, we cannot assign the glycan structures or occupancy from only the results presented in **Figures 3,4** due to the mass degeneracy and microheterogeneity of O-glycans.[40,62] To achieve in-depth glycoform analysis, we further utilized an ultrahigh-resolution 12T FTICR capable of baseline and isotopically resolving the S-RBD O-glycoforms for MS/MS analysis (see Methods for details, **Figure S4**). Following further MS$^1$ analysis, individual glycoforms were directly visualized with high resolution (**Figure 5A,B**). Following specific quadrupole isolation centered at appropriate *m/z* widths to capture individual S-RBD proteoforms (**Figure 5C**), the resolved intact glycoforms can be characterized by MS/MS using a combination of fragmentation strategies to achieve confident protein sequence assignments (**Figure 5D**).
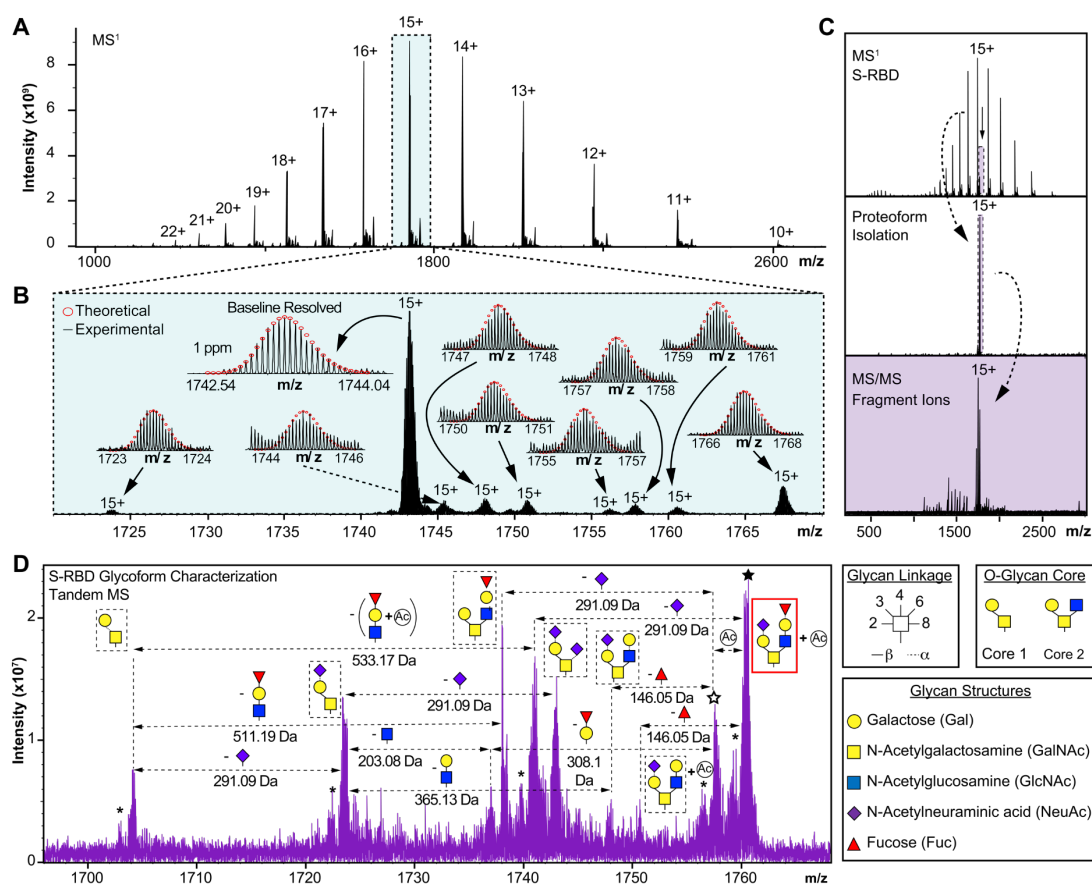
**Figure 5.** S-RBD O-glycoform proteoform analysis by high resolution top-down MS. (A) MS[1] analysis of S-RBD O-glycoforms following denaturation. (B) Magnified MS[1] from (A) showing isotopically and baseline resolved S-RBD proteoforms allowing for accurate determination of individual intact glycoforms. The most abundant charge state (15+) is isolated. Theoretical isotope distributions (red circles) are overlaid on the experimentally obtained mass spectrum to illustrate the high mass accuracy. All individual ion assignments are within 1 ppm from the theoretical mass. (C) Illustration of the isolation of a specific S-RBD glycoform using ultrahigh resolution FTICR to enable glycan site and structure characterization from intact glycoforms. (D) MS/MS characterization of S-RBD proteoform isolated from quadrupole window centered at 1760.5 *m/z*. The assignments of glycan structures are marked in the spectrum with the legends shown on the right side. Glycoform characterization reveals the specific S-RBD proteoform to have Core 2 type GalNAcGal(GalNeuAc)(GlcNAcGalFuc) glycan. Neutral loss glycan products are labeled. N-terminal acetylation (Ac) is labeled and corresponds to a +42 Da mass shift. The solid star represents the 15+ charge state precursor ion corresponding to 1760.5 *m/z*, and the hollow star represents the 15+ charge state precursor ion corresponding to 1757.7 *m/z* (-Ac). The asterisk "*" denotes an oxonium ion loss.

By applying this method on all detectable S-RBD proteoforms, a suite of MS/MS spectra were generated for sequence analysis (**Figure S5**). MS/MS data were output from the Bruker DataAnalysis software and analyzed in targeted protein analysis mode using MASH Explorer.[64] This approach allowed for the direct characterization of the glycan structures and their

microheterogeneity to reveal multiple S-RBD glycoforms with Core 1 and Core 2 O-glycan structures (**Figure 5**, and full details in **Figures S6-11**). This high-resolution method enables the structural characterization of glycoforms that have not been previously observed. As shown for one such example in **Figure 5D** (highlighted in the red box on the right), neutral loss mapping revealed a Core 2 (GlcNAcβ1-6(Galβ1-3)GalNAc-Ser/Thr) glycoform (15+ most abundant charge state, centered at 1760.5 *m/z*) with a GalNAcGal(GalNeuAc)(GlcNAcGalFuc) structure at Thr323. The site assignment of this glycan to Thr323 agrees with previous studies reporting putative assignments of other S protein O-glycans at the same site,[22] which is reasonable given the presence of an adjacent proline at residue 322, increasing the likelihood of an O-glycosylation site. Direct visualization of MS/MS fragment ions obtained from collision-induced dissociation (CID) along with the corresponding proteoform sequence mapping confirms the glycan structure assignment with high mass accuracy (**Figure S6**). Additionally, the GalNAcGal(GalNeuAc)(GlcNAcGalFuc) related proteoform without N-terminal acetylation (15+ most abundant charge state, centered at 1757.7 *m/z*) was also characterized (**Figure S6**). Therefore, the high resolution glycoform mapping presented in **Figure 5** illustrates a unique strength of this native top-down MS approach to achieve isotopically resolved and high accuracy characterization of highly heterogenous glycoproteins, a major challenge in intact glycoprotein analysis.[40,65]

We also identified and characterized core 1 (Galβ1-3GalNAc-Ser/Thr) O-glycan structures such as GalNAcGalNeuAc (15+ most abundant charge state, centered at 1723.8 *m/z*) (**Figure S7**), and GalNAcGal(NeuAc)$_2$ (15+ most abundant charge state, centered at 1743.2 *m/z*) (**Figure S8**). Interestingly, these two Core 1 O-glycans were previously reported for the S-RBD as potential modifications, however the previous studies were not able to resolve the exact glycoforms due to the challenges arising from inferring intact glycoprotein structures from peptide digests and the

signal low abundance of O-glycans under conventional MS analysis.[22,23,37] However, our native top-down MS strategy enables us to unambiguously reveal the exact glycoform corresponding to each of these Core 1 O-glycans. Two additional Core 2 glycan structures were identified GalNAcGal(GalNeuAc)(GlcNAcGal) (15+ most abundant charge state, centered at 1748.1 *m/z*, with additional N-terminal acetylated proteoform found at 1750.5 *m/z*) (**Figures S9,10**) and GalNAcGal(GalNeuAc)(GlcNAcGalNeuAc) (centered at 1767.6 *m/z*) (**Figure S11**).

This native top-down MS approach enables the proteoform-resolved characterization of S-RBD O-glycoforms to identify and characterize multiple S-RBD O-glycoforms as well as a novel fucose-containing glycoform with Core 2 GalNAcGal(GalNeuAc)(GlcNAcGalFuc) glycan structure (**Figure 6A**). **Figure 6A** shows all of the 7 identified O-glycoforms of the S-RBD, which are also listed in the table in **Figure 6B**. As a unique advantage of the top-down MS approach, the molecular abundance of each intact protein glycoform can be relatively quantified. We found that the relative abundance of Core 1 to Core 2 S-RBD O-glycan structures was roughly 70:30, with the Core 1 GalNAcGal(NeuAc)$_2$ being the most abundant O-glycoform (~67% relative abundance) (**Figure 6B**). This ability to unambiguously elucidate the structure of a specific S-RBD glycoform with high accuracy and quantify its relative abundance demonstrates the distinct advantages of this native top-down MS strategy over glycopeptide-based bottom-up MS approaches.[47,62] Together, this proteoform-resolved intact glycoprotein analysis strategy enables the simultaneous characterization of O-glycan structures of the S-RBD and its microheterogeneity, including structure and relative molecular abundance (**Figure 6C**). Importantly, the accurate determination of the relative abundance of intact glycoforms provides the technical foundation to understand the functional significance of these distinct glycoforms of S protein in the future. Although these identified O-glycoforms are specifically for S-RBD expressed from HEK293 systems, the

13

HEK293 cell expression system is known to reflect the glycosylation sites expected for the viron[21-23,27] and is currently the antigen of choice for vaccine candidates and virus functional studies.[66,67]
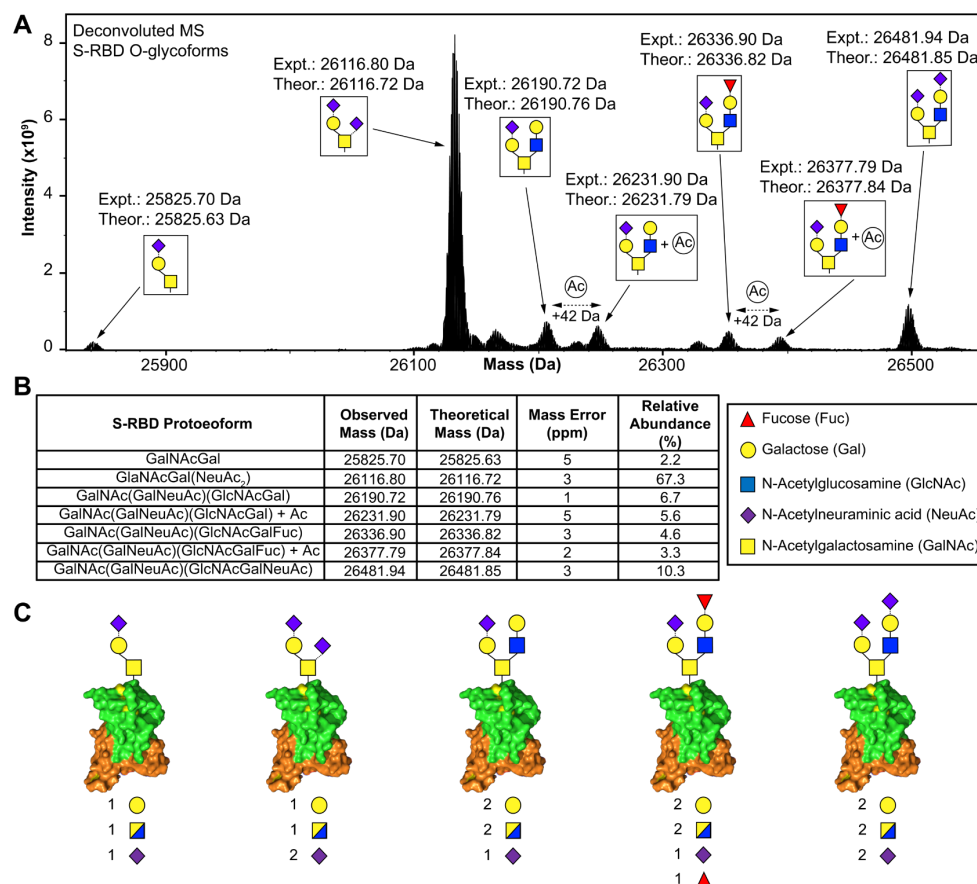


**Figure 6.** Complete characterization of the structures and relative abundance of S-RBD O-glycoforms by native top-down MS. (A) Deconvoluted MS of the S-RBD O-glycoforms showing all seven intact glycan assignments revealed by top-down MS/MS. Experimental and theoretical mass assignments are illustrated for each glycoform. (B) Summary of the O-glycoform characterization for the S-RBD with the relative abundance and mass accuracy of each identified glycoform. Masses are reported as the monoisotopic mass. (C) Illustration of S-RBD O-glycan microheterogeneity identified by specific proteoform isolation and sequencing.

## Conclusion

In summary, the combination of native trapped ion mobility MS from timsTOF and native top-down MS from ultrahigh-resolution FTICR enables the deep and comprehensive analysis of the intact glycoforms of the S-RBD, for the first time. Importantly, we report the complete structural characterization of seven S-RBD glycoforms, along with a Core 2 fucosylated glycan structure not

14

previously reported. Moreover, the relative molecular abundances of each of these intact glycoforms can be quantified, providing new insights into the complex landscape of S-RBD O-glycosylation. Our proteoform-resolved method provides complete structural characterization of O-glycan microheterogeneity and complex glycoproteins. This native top-down mass spectrometry approach not only enables the comprehensive analysis of the S-RBD and the molecular signatures of emerging S-RBD variants of SARS-CoV-2 virus toward the structure-function studies in the future, but also can help catalyze the studies of other O-glycoproteins in general.

## Materials and Methods

**Materials and Reagents.** All chemicals and reagents were purchased from MilliporeSigma (Burlington, MA, USA) and used as received without further purification unless otherwise noted. Aqueous solutions were made in nanopure deionized water ($H_2O$) from Milli-Q® water (MilliporeSigma). Ammonium acetate (AA) was purchased from Fisher Scientific (Fair Lawn, NJ). Bradford protein assay reagent was purchased from Bio-Rad (Hercules, CA, USA). 10% gel (10 or 15 comb well, 10.0 cm × 10.0 cm, 1.0 mm thick) for SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE) was home-made. Thermo Scientific™ Cimarec+™ stirring hotplate were purchased from ThermoFisher Scientific. Amicon, 0.5 mL cellulose ultra centrifugal filters with a molecular-weight cutoff (MWCO) of 10 kDa were purchased from MilliporeSigma. The recombinant SARS-CoV-2 spike regional-binding domain protein expressed in the HEK 293 cell line was purchased from Sino Biological Inc. (cat. 40592-VNAH).

**Complete N-glycan removal for O-glycan protein analysis.** Peptide-N-glycosidase F (PNGase F) (New England Biolabs Inc., Cat. P0704S) was used for complete removal of N-linked oligosaccharides. Briefly, 20 µg of protein sample was buffer exchanging into 150 mM AA (pH = 7.4) solution by washing the sample five times through a 10 kDa Amicon ultra centrifugal filters (MilliporeSigma, Burlington, MA, USA). PNGase F (5 U) was then added to the protein solution and samples were incubated at 37 °C for 12 h on a Thermo Scientific™ Cimarec+™ stirring hotplate.

**Sample preparation.** Native protein samples were prepared by buffer exchanging into 150 mM AA solution by washing the sample five times through a 10 kDa Amicon ultra centrifugal filters (MilliporeSigma, Burlington, MA, USA). The protein sample was then diluted to 20 µM in 150 mM AA prior to native top-down MS analysis. Denatured protein samples were reduced using 20 mM tris(2-carboxyethyl)phosphine hydrochloride (TCEP), and followed by buffer exchanging into 0.1 % formic acid solution by washing the sample five times through a 10 kDa Amicon ultra centrifugal filters (MilliporeSigma, Burlington, MA, USA). The protein sample was then diluted to 20 µM in 50:50:0.1 (acetonitrile: water: formic acid) prior to top-down MS analysis. Protein samples subjected to N-glycan removal were prepared similarly following PNGase F reaction.

**Top-down MS analysis.** Intact protein samples were analyzed by nano-electrospray ionization via direct infusion using a TriVersa Nanomate system (Advion BioSciences, Ithaca, NY, USA) coupled to a solariX XR 12-Tesla Fourier Transform Ion Cyclotron Resonance mass spectrometer (FTICR-MS, Bruker Daltonics, Billerica, MA). For the nano-electrospray ionization source using a TriVersa Nanomate, the desolvating gas pressure was set at 0.45 PSI and the voltage was set to

16

1.2 to 1.6 kV versus the inlet of the mass spectrometer. Source dry gas flow rate was set to 4 L/min at 124 °C. For the source optics, the capillary exit, deflector plate, funnel 1, skimmer voltage, funnel RF amplitude, octopole frequency, octopole RF amplitude, collision cell RF frequency, and collision cell RF amplitude were optimized at 190 V, 200 V, 100 V, 35 V, 250 Vpp, 2 MHz, 490 Vpp, 2MHz, and 2000 Vpp, respectively. Mass spectra were acquired with an acquisition size of 4M, in the mass range between 200-4000 *m/z* (with a resolution of 530,000 at 400 *m/z*), and a minimum of 250 scans were accumulated for each sample. Ions were accumulated in the collision cell for 0.100 s, and a time of flight of 1.500 ms was used prior to their transfer to the ICR cell. For collision-induced dissociation (CID) tandem MS (MS/MS) experiments, the collision energy was varied from 10 to 30V, ion accumulation was optimized at 400 ms, and file size was 4M. Tandem mass spectra were output from the DataAnalysis 4.3 (Bruker Daltonics) software and analyzed using MASH Explorer.[64]

**Native top-down MS analysis.** A timsTOF Pro mass spectrometer (Bruker Daltonics, Bremen, Germany) was couple to a Bruker nanoElute LC system (Bruker Daltonics, Bremen, Germany). Samples were directly infused using the nanoElute, injecting 5 μL of native protein sample with a flow rate of 1 μL/min. For the MS inlet, the endplate offset and capillary voltage were set to 500 V and 3800 V, respectively. The nebulizer gas pressure ($N_2$) was set to 1.5 bar, with a dry gas flow rate of 6 L/min at 180 °C. The tunnel out, tunnel in, and TOF vacuum pressures were set to 0.8584 mBar, 2.577 mBar, and 1.752E-07 mBar. To calibrate the MS and trapped ion mobility spectrometry (TIMS) device, Agilent tune mix was directly infused to provide species of known mass and reduced mobility.[37] For MS calibration, the MS resolution for the most abundant calibrant signal, 1222 *m/z*, was 62,000. Calibrant points at 922, 1222, and 1522 *m/z* were used for

TIMS calibration. The TIMS resolution for the most abundant calibrant signal, 1222 *m/z*, was 80.2 CCS/ΔCCS. IMS tunnel voltage deltas were optimized at -20 V, -120 V, 70 V, 200 V, 0 V and 100 V for Δ1 to Δ6, respectively. TIMS funnel 1 RF was set to 350 Vpp, and TIMS collision cell energy was set to 250 V. An IMS imeX accumulation time and cycle ramp time of 400 ms was found to yield optimal TIMS resolving power. To facilitate high MS signal intensity, the TIMS accumulation time was locked to a 100% duty cycle, or 400 ms accumulation time. In the MS transfer optics the funnel 1 RF, funnel 2 RF, deflection delta, in-source collision induced dissociation (isCID) energy, multipole RF, and quadrupole ion energy were optimized at, 350 Vpp, 500 Vpp, 55 V, 0 eV, 800 Vpp, and 4 eV, respectively. For MS1 spectral collection, the quadrupole low mass was set to 1000 *m/z* with a scan range of 200 to 5000 *m/z*. Collision energy was set to 6 eV, with 3500 Vpp collision cell RF, a 100 μs transfer time, and a pre pulse storage time of 35 μs. For MS2 spectral collection the quadrupole low mass was set to 200 *m/z* with a scan range of 200 to 5000 *m/z*. The collision cell RF was set to 3500 Vpp, a 100 μs transfer time, and a pre pulse storage time of 20 μs were used. Mass spectra were output from the DataAnalysis 5.3 (Bruker Daltonics) software and analyzed using MASH Explorer.[64]

**Data Analysis.** All data were processed and analyzed using Compass DataAnalysis 4.3/5.3 and MASH Explorer.[64] Maximum Entropy algorithm (Bruker Daltonics) was used to deconvolute all mass spectra with the resolution set to 80,000 for the timsTOF Pro or with instrument peak width set to 0.05 for the 12T FTICR. Sophisticated Numerical Annotation Procedure (SNAP) peak-picking algorithm (quality factor: 0.4; signal-to-noise ratio (S/N): 3.0; intensity threshold: 500) was applied to determine the monoisotopic mass of all detected ions. The relative abundance for each protein isoform was determined using DataAnalysis. To quantify protein modifications, the

18

relative abundances of specific modifications were calculated as their corresponding percentages among all the detected protein forms in the deconvoluted averaged mass. MS/MS data were output from the DataAnalysis software and analyzed using MASH Explorer[64] for proteoform identification and sequence mapping. All the program-processed data were manually validated. For peak picking, the sophisticated numerical annotation procedure (SNAP) algorithm from Bruker DataAnalysis 5.3 was used with a quality threshold of 0.5 and an S/N lower threshold of 3. All fragment ions were manually validated using MASH Explorer. Peak extraction was performed using a signal-to-noise ratio of 3 and a minimum fit of 60%, and all peaks were subjected to manual validation. All identifications were made with satisfactory numbers of assigned fragment (>10), and a 25-ppm mass tolerance was used to match the experimental fragment ions to the calculated fragment ions based on amino acid sequence. For ion mobility analysis by the timsTOF Pro, the collisional cross section (CCS) in $\text{Å}^2$ for a species of interest was calculated via the Mason-Schamps equation (Equation 1):

$$CCS = \frac{3}{16}\sqrt{\frac{2\pi}{\mu k_b T}}\frac{ze}{N_0 K_0}$$

(1)

Where $\mu$ is the reduced mass of the ion-gas pair ($\mu = \frac{mM}{(m+M)}$, where m and M are the ion and gas particle masses), $k_b$ is Boltzmann's constant, $T$ is the drift region temperature, $z$ is the ionic charge, $e$ is the charge of an electron, $N_0$ is the buffer gas density, and $K_0$ is the reduced mobility. Equation 1 was selected to agree with previously published CCS calculations.[57,58,68] Theoretical CCS were calculated using the IMPACT method.[60]

**Supporting Information**

The Supporting Information is available free of charge on the ACS Publications website.

19

## Acknowledgement

## Competing interests

The authors declare no competing financial interest.

## References

(1) Wu, F.; Zhao, S.; Yu, B.; Chen, Y.-M.; Wang, W.; Song, Z.-G.; Hu, Y.; Tao, Z.-W.; Tian, J.-H.; Pei, Y.-Y.et al. A new coronavirus associated with human respiratory disease in China. *Nature* **2020,** *579* (7798), 265.

(2) World Health Organization. *Coronavirus Disease (COVID-19) Weekly Epidemiological Update for February 2nd 2021; WHO*, 2021.

(3) Walls, A. C.; Park, Y.-J.; Tortorici, M. A.; Wall, A.; McGuire, A. T.; Veesler, D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **2020**.

(4) Yan, R.; Zhang, Y.; Li, Y.; Xia, L.; Guo, Y.; Zhou, Q. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* **2020,** *367* (6485), 1444.

(5) Liu, C.; Zhou, Q.; Li, Y.; Garner, L. V.; Watkins, S. P.; Carter, L. J.; Smoot, J.; Gregg, A. C.; Daniels, A. D.; Jervey, S.et al. Research and Development on Therapeutic Agents and Vaccines for COVID-19 and Related Human Coronavirus Diseases. *ACS Central Science* **2020**.

(6) Folegatti, P. M.; Ewer, K. J.; Aley, P. K.; Angus, B.; Becker, S.; Belij-Rammerstorfer, S.; Bellamy, D.; Bibi, S.; Bittaye, M.; Clutterbuck, E. A.et al. Safety and immunogenicity of the ChAdOx1 nCoV-19 vaccine against SARS-CoV-2: a preliminary report of a phase 1/2, single-blind, randomised controlled trial. *The Lancet* **2020,** *396* (10249), 467.

(7) Jackson, L. A.; Anderson, E. J.; Rouphael, N. G.; Roberts, P. C.; Makhene, M.; Coler, R. N.; McCullough, M. P.; Chappell, J. D.; Denison, M. R.; Stevens, L. J.et al. An mRNA Vaccine against SARS-CoV-2 — Preliminary Report. *New England Journal of Medicine* **2020,** *383* (20), 1920.

(8) Chen, P.; Nirula, A.; Heller, B.; Gottlieb, R. L.; Boscia, J.; Morris, J.; Huhn, G.; Cardona, J.; Mocherla, B.; Stosor, V.et al. SARS-CoV-2 Neutralizing Antibody LY-CoV555 in Outpatients with Covid-19. *New England Journal of Medicine* **2020,** *384* (3), 229.

(9) Hansen, J.; Baum, A.; Pascal, K. E.; Russo, V.; Giordano, S.; Wloga, E.; Fulton, B. O.; Yan, Y.; Koon, K.; Patel, K.et al. Studies in humanized mice and convalescent humans yield a SARS-CoV-2 antibody cocktail. *Science* **2020,** *369* (6506), 1010.

(10) Starr, T. N.; Greaney, A. J.; Hilton, S. K.; Ellis, D.; Crawford, K. H. D.; Dingens, A. S.; Navarro, M. J.; Bowen, J. E.; Tortorici, M. A.; Walls, A. C.et al. Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. *Cell* **2020,** *182* (5), 1295.

(11) Shu, Y.; McCauley, J. GISAID: Global initiative on sharing all influenza data – from vision to reality. *Eurosurveillance* **2017,** *22* (13), 30494.

(12) Meredith, L. W.; Hamilton, W. L.; Warne, B.; Houldcroft, C. J.; Hosmillo, M.; Jahun, A. S.; Curran, M. D.; Parmar, S.; Caller, L. G.; Caddy, S. L.et al. Rapid implementation of SARS-CoV-2 sequencing to investigate cases of health-care associated COVID-19: a prospective genomic surveillance study. *The Lancet Infectious Diseases* **2020,** *20* (11), 1263.

(13) Thomson, E. C.; Rosen, L. E.; Shepherd, J. G.; Spreafico, R.; da Silva Filipe, A.; Wojcechowskyj, J. A.; Davis, C.; Piccoli, L.; Pascall, D. J.; Dillen, J.et al. Circulating SARS-CoV-2 spike N439K variants maintain fitness while evading antibody-mediated immunity. *Cell* **2021**.

(14) Korber, B.; Fischer, W. M.; Gnanakaran, S.; Yoon, H.; Theiler, J.; Abfalterer, W.; Hengartner, N.; Giorgi, E. E.; Bhattacharya, T.; Foley, B.et al. Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. *Cell* **2020,** *182* (4), 812.

(15) Walls, A. C.; Xiong, X.; Park, Y.-J.; Tortorici, M. A.; Snijder, J.; Quispe, J.; Cameroni, E.; Gopal, R.; Dai, M.; Lanzavecchia, A.et al. Unexpected Receptor Functional Mimicry Elucidates Activation of Coronavirus Fusion. *Cell* **2019,** *176* (5), 1026.

(16) Andersen, K. G.; Rambaut, A.; Lipkin, W. I.; Holmes, E. C.; Garry, R. F. The proximal origin of SARS-CoV-2. *Nature Medicine* **2020,** *26* (4), 450.

(17) Casalino, L.; Gaieb, Z.; Goldsmith, J. A.; Hjorth, C. K.; Dommer, A. C.; Harbison, A. M.; Fogarty, C. A.; Barros, E. P.; Taylor, B. C.; McLellan, J. S.et al. Beyond Shielding: The Roles of Glycans in the SARS-CoV-2 Spike Protein. *ACS Central Science* **2020,** *6* (10), 1722.

(18) Xiong, X.; Tortorici, M. A.; Snijder, J.; Yoshioka, C.; Walls, A. C.; Li, W.; McGuire, A. T.; Rey, F. A.; Bosch, B.-J.; Veesler, D. Glycan Shield and Fusion Activation of a Deltacoronavirus Spike Glycoprotein Fine-Tuned for Enteric Infections. *Journal of Virology* **2018,** *92* (4), e01628.

(19) Walls, A. C.; Tortorici, M. A.; Frenz, B.; Snijder, J.; Li, W.; Rey, F. A.; DiMaio, F.; Bosch, B.-J.; Veesler, D. Glycan shield and epitope masking of a coronavirus spike protein observed by cryo-electron microscopy. *Nature Structural & Molecular Biology* **2016,** *23* (10), 899.

(20) Grant, O. C.; Montgomery, D.; Ito, K.; Woods, R. J. Analysis of the SARS-CoV-2 spike protein glycan shield reveals implications for immune recognition. *Scientific Reports* **2020,** *10* (1), 14991.

(21) Watanabe, Y.; Allen, J. D.; Wrapp, D.; McLellan, J. S.; Crispin, M. Site-specific glycan analysis of the SARS-CoV-2 spike. *Science* **2020**, eabb9983.

(22) 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**.

(23) Sanda, M.; Morrison, L.; Goldman, R. N- and O-Glycosylation of the SARS-CoV-2 Spike Protein. *Analytical Chemistry* **2021,** *93* (4), 2003.

(24) Jensen, P. H.; Kolarich, D.; Packer, N. H. Mucin-type O-glycosylation – putting the pieces together. *The FEBS Journal* **2010,** *277* (1), 81.

(25) Darula, Z.; Medzihradszky, K. F. Analysis of Mammalian O-Glycopeptides—We Have Made a Good Start, but There is a Long Way to Go*. *Molecular & Cellular Proteomics* **2018,** *17* (1), 2.

(26) Varki, A.; Cummings, R. D.; Esko, J. D.; Stanley, P.; Hart, G. W.; Aebi, M.; Darvill, A. G.; Kinoshita, T.; Packer, N. H.; Prestegard, J. H. *Essentials of Glycobiology [Internet]*; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2015.

(27) Bagdonaite, I.; Wandall, H. H. Global aspects of viral glycosylation. *Glycobiology* **2018,** *28* (7), 443.

(28) Varki, A. Biological roles of glycans. *Glycobiology* **2016,** *27* (1), 3.

(29)  Francés-Monerris, A.; Hognon, C.; Miclot, T.; García-Iriepa, C.; Iriepa, I.; Terenzi, A.; Grandemange, S.; Barone, G.; Marazzi, M.; Monari, A. Molecular Basis of SARS-CoV-2 Infection and Rational Design of Potential Antiviral Agents: Modeling and Simulation Approaches. *Journal of Proteome Research* **2020**.

(30)  Wu, D.; Struwe, W. B.; Harvey, D. J.; Ferguson, M. A. J.; Robinson, C. V. N-glycan microheterogeneity regulates interactions of plasma proteins. *Proceedings of the National Academy of Sciences* **2018**, *115* (35), 8763.

(31)  Riley, N. M.; Hebert, A. S.; Westphall, M. S.; Coon, J. J. Capturing site-specific heterogeneity with large-scale N-glycoproteome analysis. *Nature Communications* **2019**, *10* (1), 1311.

(32)  Khatri, K.; Klein, J. A.; White, M. R.; Grant, O. C.; Leymarie, N.; Woods, R. J.; Hartshorn, K. L.; Zaia, J. Integrated Omics and Computational Glycobiology Reveal Structural Basis for Influenza A Virus Glycan Microheterogeneity and Host Interactions*. *Molecular & Cellular Proteomics* **2016**, *15* (6), 1895.

(33)  Serapian, S. A.; Marchetti, F.; Triveri, A.; Morra, G.; Meli, M.; Moroni, E.; Sautto, G. A.; Rasola, A.; Colombo, G. The Answer Lies in the Energy: How Simple Atomistic Molecular Dynamics Simulations May Hold the Key to Epitope Prediction on the Fully Glycosylated SARS-CoV-2 Spike Protein. *The Journal of Physical Chemistry Letters* **2020**, *11* (19), 8084.

(34)  Mayr, J.; Lau, K.; Lai, J. C. C.; Gagarinov, I. A.; Shi, Y.; McAtamney, S.; Chan, R. W. Y.; Nicholls, J.; von Itzstein, M.; Haselhorst, T. Unravelling the Role of O-glycans in Influenza A Virus Infection. *Scientific Reports* **2018**, *8* (1), 16382.

(35)  Suenaga, T.; Matsumoto, M.; Arisawa, F.; Kohyama, M.; Hirayasu, K.; Mori, Y.; Arase, H. Sialic Acids on Varicella-Zoster Virus Glycoprotein B Are Required for Cell-Cell Fusion. *Journal of Biological Chemistry* **2015**, *290* (32), 19833.

(36)  Falkowska, E.; Kajumo, F.; Garcia, E.; Reinus, J.; Dragic, T. Hepatitis C Virus Envelope Glycoprotein E2 Glycans Modulate Entry, CD81 Binding, and Neutralization. *Journal of Virology* **2007**, *81* (15), 8072.

(37)  Riley, N. M.; Bertozzi, C. R.; Pitteri, S. J. A Pragmatic Guide to Enrichment Strategies for Mass Spectrometry–Based Glycoproteomics. *Molecular & Cellular Proteomics* **2021**, *20*, 100029.

(38)  Wilkinson, H.; Saldova, R. Current Methods for the Characterization of O-Glycans. *Journal of Proteome Research* **2020**, *19* (10), 3890.

(39)  Smith, L. M.; Kelleher, N. L.; The Consortium for Top Down, P.; Linial, M.; Goodlett, D.; Langridge-Smith, P.; Ah Goo, Y.; Safford, G.; Bonilla*, L.; Kruppa, G.et al. Proteoform: a single term describing protein complexity. *Nature Methods* **2013**, *10*, 186.

(40)  Wang, G.; de Jong, R. N.; van den Bremer, E. T. J.; Parren, P. W. H. I.; Heck, A. J. R. Enhancing Accuracy in Molecular Weight Determination of Highly Heterogeneously Glycosylated Proteins by Native Tandem Mass Spectrometry. *Analytical Chemistry* **2017**, *89* (9), 4793.

(41)  Leney, A. C.; Rafie, K.; van Aalten, D. M. F.; Heck, A. J. R. Direct Monitoring of Protein O-GlcNAcylation by High-Resolution Native Mass Spectrometry. *ACS Chemical Biology* **2017**, *12* (8), 2078.

(42)  Yang, Y.; Liu, F.; Franc, V.; Halim, L. A.; Schellekens, H.; Heck, A. J. R. Hybrid mass spectrometry approaches in glycoprotein analysis and their usage in scoring biosimilarity. *Nature Communications* **2016**, *7* (1), 13397.

(43)  Tiambeng, T. N.; Roberts, D. S.; Brown, K. A.; Zhu, Y.; Chen, B.; Wu, Z.; Mitchell, S. D.; Guardado-Alvarez, T. M.; Jin, S.; Ge, Y. Nanoproteomics enables proteoform-resolved analysis of low-abundance proteins in human serum. *Nature Communications* **2020**, *11* (1), 3903.

(44)  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* **2020**, null.

(45)  Chen, B.; Brown, K. A.; Lin, Z.; Ge, Y. Top-Down Proteomics: Ready for Prime Time? *Analytical Chemistry* **2018**, *90* (1), 110.

22

(46)   Vimer, S.; Ben-Nissan, G.; Morgenstern, D.; Kumar-Deshmukh, F.; Polkinghorn, C.; Quintyn, R. S.; Vasil'ev, Y. V.; Beckman, J. S.; Elad, N.; Wysocki, V. H.et al. Comparative Structural Analysis of 20S Proteasome Ortholog Protein Complexes by Native Mass Spectrometry. *ACS Central Science* **2020,** *6* (4), 573.

(47)   Zhou, M.; Lantz, C.; Brown, K. A.; Ge, Y.; Paša-Tolić, L.; Loo, J. A.; Lermyte, F. Higher-order structural characterisation of native proteins and complexes by top-down mass spectrometry. *Chemical Science* **2020**.

(48)   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* **2018,** *10*, 139.

(49)   Wohlschlager, T.; Scheffler, K.; Forstenlehner, I. C.; Skala, W.; Senn, S.; Damoc, E.; Holzmann, J.; Huber, C. G. Native mass spectrometry combined with enzymatic dissection unravels glycoform heterogeneity of biopharmaceuticals. *Nature Communications* **2018,** *9* (1), 1713.

(50)   Rabuck-Gibbons, J. N.; Keating, J. E.; Ruotolo, B. T. Collision induced unfolding and dissociation differentiates ATP-competitive from allosteric protein tyrosine kinase inhibitors. *International Journal of Mass Spectrometry* **2018,** *427*, 151.

(51)   Leney, A. C.; Heck, A. J. R. Native Mass Spectrometry: What is in the Name? *Journal of the American Society for Mass Spectrometry* **2017,** *28* (1), 5.

(52)   Hernandez, H.; Robinson, C. V. Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. *Nat Protoc* **2007,** *2* (3), 715.

(53)   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* **2020,** *92* (6), 4459.

(54)   Wei, J.; Tang, Y.; Ridgeway, M. E.; Park, M. A.; Costello, C. E.; Lin, C. Accurate Identification of Isomeric Glycans by Trapped Ion Mobility Spectrometry-Electronic Excitation Dissociation Tandem Mass Spectrometry. *Analytical Chemistry* **2020,** *92* (19), 13211.

(55)   Michelmann, K.; Silveira, J. A.; Ridgeway, M. E.; Park, M. A. Fundamentals of Trapped Ion Mobility Spectrometry. *Journal of the American Society for Mass Spectrometry* **2015,** *26* (1), 14.

(56)   Jeanne Dit Fouque, K.; Garabedian, A.; Leng, F.; Tse-Dinh, Y.-C.; Ridgeway, M. E.; Park, M. A.; Fernandez-Lima, F. Trapped Ion Mobility Spectrometry of Native Macromolecular Assemblies. *Analytical Chemistry* **2021,** *93* (5), 2933.

(57)   Dodds, J. N.; Baker, E. S. Ion Mobility Spectrometry: Fundamental Concepts, Instrumentation, Applications, and the Road Ahead. *Journal of the American Society for Mass Spectrometry* **2019,** *30* (11), 2185.

(58)   Kirk, S. R.; Liu, F. C.; Cropley, T. C.; Carlock, H. R.; Bleiholder, C. On the Preservation of Non-covalent Peptide Assemblies in a Tandem-Trapped Ion Mobility Spectrometer-Mass Spectrometer (TIMS-TIMS-MS). *Journal of the American Society for Mass Spectrometry* **2019,** *30* (7), 1204.

(59)   Benigni, P.; Marin, R.; Molano-Arevalo, J. C.; Garabedian, A.; Wolff, J. J.; Ridgeway, M. E.; Park, M. A.; Fernandez-Lima, F. Towards the analysis of high molecular weight proteins and protein complexes using TIMS-MS. *International Journal for Ion Mobility Spectrometry* **2016,** *19* (2), 95.

(60)   Marklund, Erik G.; Degiacomi, Matteo T.; Robinson, Carol V.; Baldwin, Andrew J.; Benesch, Justin L. P. Collision Cross Sections for Structural Proteomics. *Structure* **2015,** *23* (4), 791.

(61)   Lan, J.; Ge, J.; Yu, J.; Shan, S.; Zhou, H.; Fan, S.; Zhang, Q.; Shi, X.; Wang, Q.; Zhang, L.et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **2020,** *581* (7807), 215.

(62)   Yen, H.-Y.; Liko, I.; Gault, J.; Wu, D.; Struwe, W. B.; Robinson, C. V. Correlating Glycoforms of DC-SIGN with Stability Using a Combination of Enzymatic Digestion and Ion Mobility Mass Spectrometry. *Angewandte Chemie International Edition* **2020,** *59* (36), 15560.

(63)    Kuprowski, M. C.; Konermann, L. Signal Response of Coexisting Protein Conformers in Electrospray Mass Spectrometry. *Analytical Chemistry* **2007,** *79* (6), 2499.

(64)    Wu, Z.; Roberts, D. S.; Melby, J. A.; Wenger, K.; Wetzel, M.; Gu, Y.; Ramanathan, S. G.; Bayne, E. F.; Liu, X.; Sun, R.et al. MASH Explorer: A Universal Software Environment for Top-Down Proteomics. *Journal of Proteome Research* **2020,** *19* (9), 3867.

(65)    McIlwain, S. J.; Wu, Z.; Wetzel, M.; Belongia, D.; Jin, Y.; Wenger, K.; Ong, I. M.; Ge, Y. Enhancing Top-Down Proteomics Data Analysis by Combining Deconvolution Results through a Machine Learning Strategy. *Journal of the American Society for Mass Spectrometry* **2020**.

(66)    Yang, J.; Wang, W.; Chen, Z.; Lu, S.; Yang, F.; Bi, Z.; Bao, L.; Mo, F.; Li, X.; Huang, Y.et al. A vaccine targeting the RBD of the S protein of SARS-CoV-2 induces protective immunity. *Nature* **2020,** *586* (7830), 572.

(67)    Shaan Lakshmanappa, Y.; Elizaldi, S. R.; Roh, J. W.; Schmidt, B. A.; Carroll, T. D.; Weaver, K. D.; Smith, J. C.; Verma, A.; Deere, J. D.; Dutra, J.et al. SARS-CoV-2 induces robust germinal center CD4 T follicular helper cell responses in rhesus macaques. *Nature Communications* **2021,** *12* (1), 541.

(68)    Revercomb, H. E.; Mason, E. A. Theory of plasma chromatography/gaseous electrophoresis. Review. *Analytical Chemistry* **1975,** *47* (7), 970.