Journal of proteome research

pubs.acs.org/jpr

Article

Direct Comparison of *N*-Glycans and Their Isomers Derived from Spike Glycoprotein 1 of MERS-CoV, SARS-CoV-1, and SARS-CoV-2

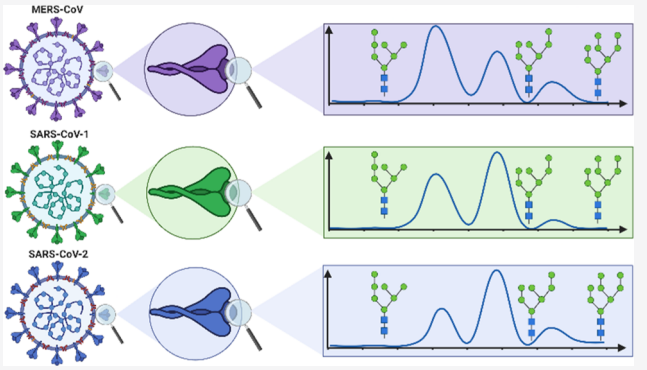
Byeong Gwan Cho, Sakshi Gautam, Wenjing Peng, Yifan Huang, Mona Goli, and Yehia Mechref*

Cite This: https://doi.org/10.1021/acs.jproteome.1c00323

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The emergence of COVID-19 pandemic has engaged the scientific community around the globe in the rapid development of effective therapeutics and vaccines. Owing to its crucial role in the invasion of the host cell, spike (S) glycoprotein is one of the major targets in these studies. The S1 subunit of the S protein (S1 protein) accommodates the receptor-binding domain, which enables the initial binding of the virus to the host cell. Being a heavily glycosylated protein, numerous studies have investigated its glycan composition. However, none of the studies have explored the isomeric glycan distribution of this protein. Furthermore, this isomeric glycan distribution has never been compared to that in S1 proteins of other coronaviruses, severe acute respiratory syndrome coronavirus 1 and Middle East respiratory syndrome coronavirus, which were responsible for past epidemics. This study explores the uncharted territory of the isomeric glycan distribution in the coronaviruses' S1 protein using liquid chromatography coupled to tandem mass spectrometry. We believe that our data would facilitate future investigations to study the role of isomeric glycans in coronavirus viral pathogenesis.



KEYWORDS: *permethylated glycan, SARS-CoV-2, LC−MS/MS, isomeric separation, RPLC*

INTRODUCTION

The novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), believed to have originated in China's Hubei province in December 2019, was identified as the causative agent of rapidly spreading Coronavirus Disease 2019 (COVID-19).[1] The disease was declared a global pandemic by the World Health Organization (WHO) in March 2020. As of April 7, 2021, the number of reported cases worldwide has risen to over 130 million with over 2.5 million reported deaths according to the WHO. However, it is plausible that many cases and related deaths go unreported due to a lack of adequate testing in some parts of the world.[2,3] Efforts are ongoing around the world with researchers persistently working to gather scientific data about the virus and develop effective treatments and vaccines.[4−6]

SARS-CoV-2 belongs to the coronavirus family which includes viruses, such as severe acute respiratory syndrome coronavirus 1 (SARS-CoV-1) and Middle East respiratory syndrome coronavirus (MERS-CoV), that are also known to cause severe respiratory infections in humans.[7,8] The hallmark of SARS-CoV-2 pathogenesis is the viral transmembrane spike (S) glycoprotein that is instrumental in attacking the host cells by interacting with the angiotensin-converting enzyme 2 receptor through its receptor-binding domain (RBD). The S protein, a trimeric class I fusion protein, consists of two subunits: S1 and S2. The S1 subunit, which includes the RBD, is responsible for the initial binding of the virus to the host cell receptor, while S2 facilitates the fusion of viral and host cellular membranes.[9−14]

Attributable to its pivotal role in the pathogenesis, the S protein is a major target for developing therapeutics and vaccines for COVID-19. RBD of the S1 subunit is demonstrated to be targeted by the neutralizing antibodies in response to coronaviruses.[9] The S1 subunit is known to be heavily glycosylated, a state that protects the SARS-CoV-2 virus by acting as a sort of "glycan shield",[10,11] a significant structural feature that plays a crucial role in the overall viral pathogenesis and shielding of the virus from the host immune response. Thus, this viral adaptation has prompted extensive investigations of the glycosylation of the SARS-CoV-2 S1 protein.[12,13,15,16] However, there has never been a direct comparison of glycosylation between previous epidemic coronaviruses (SARS-CoV-1 and MERS-CoV) as well as

Received: April 19, 2021

glycan isomers of their glycan shield. Glycan isomers are important features of viral transmissions, and their functions have been well documented previously with flu virus variants.[17] Glycan isomers have been reported to contribute to the virion and the host interaction. One example of glycan isomers being associated with the viral interaction of the host is the influenza virus.[17] Sialic acid linkage isomers have been correlated with the effect of the influenza virus infection that varies depending on the host. A remarkable example is the human influenza virus, which is known for switching to the α-2,6 sialic acid linkage-based interaction from α-2,3 sialic acid linkage during adaptation from animals to humans. This variation enables the virus to attach to the ciliated human epithelial target cells that express α-2,6-linked sialic acids. Moreover, it escapes the α-2,3-linked sialic acids expressed on secreted soluble airway mucins that inhibit virus binding.[18] Coronaviruses are also reported to interact with the sialoglycans present on the surface of the target cells.[19]

To investigate these glycans, liquid chromatography coupled to tandem mass spectrometry (LC−MS/MS) is commonly utilized. Glycans are often chemically labeled or derivatized to enhance their separation and analysis.[20,21] One of the major derivatization techniques to stabilize glycans is permethylation.[20,22,23] Permethylation offers several advantages for glycan analysis. This chemical modification improves the ionization efficiency of the glycans. Also, it prevents sialic acid loss and inhibits fucose migration, thus stabilizing the glycans. Moreover, it enhances the glycans' hydrophobicity and facilitates their separation using reversed-phase liquid chromatography (RPLC).[20] These advantages prompted the utilization of permethylation as a method of choice for derivatizing the glycans in the current study.

Here, in this work, we demonstrate the direct comparison of glycosylation between the current pandemic coronavirus and the other two previous epidemic coronaviruses as well as the isomers of their glycans using LC−MS/MS. Moreover, we believe that in the future, it would be interesting to further investigate the role of sialic acid linkage-based isomers in SARS-CoV-2 pathogenesis to see if any relationship between the type of the host and viral sialic acid linkage preferences, as found in human influenza virus, exists.

METHODS

S1 Protein Acquisition

SARS-CoV-2 S1 and MERS-CoV S1 proteins expressed in human embryonic kidney 293 (HEK293) cells were acquired from Sino Biological (40591-V08H and 40069-V08H). SARS-CoV-1 S1 protein was purchased from Acro Biosystems (S1N-S52H5).

Sample Validation Proteomic Experiment

SARS-CoV S1 protein, SARS-CoV-2 S1 protein, and MERS S1 protein were dissolved in 50 mM ammonium bicarbonate (Sigma-Aldrich) buffer. Then, samples were reduced by 5 mM dithiothreitol (DTT) (Sigma-Aldrich) at 60 °C for 45 min followed by 20 mM iodoacetamide (IAA) (Sigma-Aldrich) alkylation at 37 °C in the dark for 45 min. The IAA reaction was then quenched by adding 5 mM DTT and incubating at 37 °C for 30 min. Next, trypsin/Lyc-C (Promega) was added to the samples with a 1:25 enzyme to protein ratio and incubated at 37 °C for 18 h. After tryptic digestion, the samples were vacuum-dried and resuspended in the loading solvent (2% ACN, 98% water, and 0.1% formic acid). The samples were prepared in three replicates.

Sequential Filter-Aided *N*-Glycomics

5 μg of each S1 protein samples originated from the three different coronaviruses (SARS-CoV-2, SARS-CoV-1, and MERS-CoV) were diluted to 50 μL of HPLC-grade water (Avantor Performance Materials) and denatured at a 90 °C water bath for 20 min. 10k MWCO filter devices (MilliporeSigma) were washed with 0.5 mL of HPLC-grade water by centrifuging (Sorvall Legend Micro 21 Centrifuge, Thermo Scientific) at 14k*g* for 20 min. Denatured S1 samples were added to the filter devices and centrifuged again at the same speed for 20 min. Filter devices were washed twice with 100 μL of Glycobuffer1 1× and centrifuged. Flow-throughs were discarded. 45 μL of Glycobuffer1 1× and 5 μL of exoglycosidase (either α2-3 neuraminidase S, 40 units or α1-3,4 fucosidase, 20 units, both from New England Biolabs) were added to the filter devices, the caps were closed, and then incubated at 37 °C water bath for 18 h. After incubation, filter devices were centrifuged at 14k*g* for 20 min. Filter devices were washed with 100 μL of 50 mM ammonium bicarbonate buffer twice by centrifuging at 14k*g* for 20 min. Flow-throughs were discarded, 49 μL of 50 mM ammonium bicarbonate buffer and 1 μL (500 units) of PNGase F (New England Biolabs) solutions were added, the caps were closed, and then incubated at 37 °C water bath for 18 h. After the incubation, exoglycosidase digested *N*-glycans were collected by centrifuging the filter devices. An additional 100 μL of 50 mM ammonium bicarbonate buffer was added twice, centrifuged, and flow throughs were collected. Samples without exoglycosidase treatment followed the same steps from above except for the exoglycosidase steps. Released glycans were dried under vacuum (Labconco CentriVap Benchtop Vacuum Concentrator).

Glycan Reduction and Permethylation

Reducing the released glycans was done following the previously reported protocol.[24] Briefly, samples were incubated in a 60 °C water bath for 1 h after being dissolved in 10 μL of fresh borane−ammonia (Sigma-Aldrich) aqueous solution (10 μg/μL). The remaining borane was then removed from the samples in the form of methyl borate by adding 0.5 mL of methanol (Fisher Scientific) and drying under vacuum with a vacuum concentrator. Methanol washing and drying were repeated several times until the borane was completely removed from the samples and no white residue remained. After the last drying, reduced glycans were subjected to solid-phase permethylation using the previously published protocol.[25] Sodium hydroxide beads (Sigma-Aldrich) were soaked in dimethyl sulfoxide (DMSO, Fisher Scientific), packed in empty spin columns (Harvard Apparatus), and washed twice with 200 μL of DMSO by centrifuging at 1800 rpm for 2 min. Reduced glycans were dissolved in 30 μL of DMSO, 1.2 μL of water, and 20 μL of iodomethane and then loaded into the sodium hydroxide bead-filled columns. The reaction mixtures were incubated at room temperature for 25 min. Afterward, an additional 20 μL of iodomethane was added to each column and incubated for 15 min. After incubation, the permethylated samples were spun down and collected using a centrifuge at 1800 rpm for 2 min. The columns were then washed with 30 μL of acetonitrile (MeCN) to elute all the remaining samples. Finally, the permethylated samples were dried using the vacuum concentrator.

https://doi.org/10.1021/acs.jproteome.1c00323
J. Proteome Res. XXXX, XXX, XXX−XXX

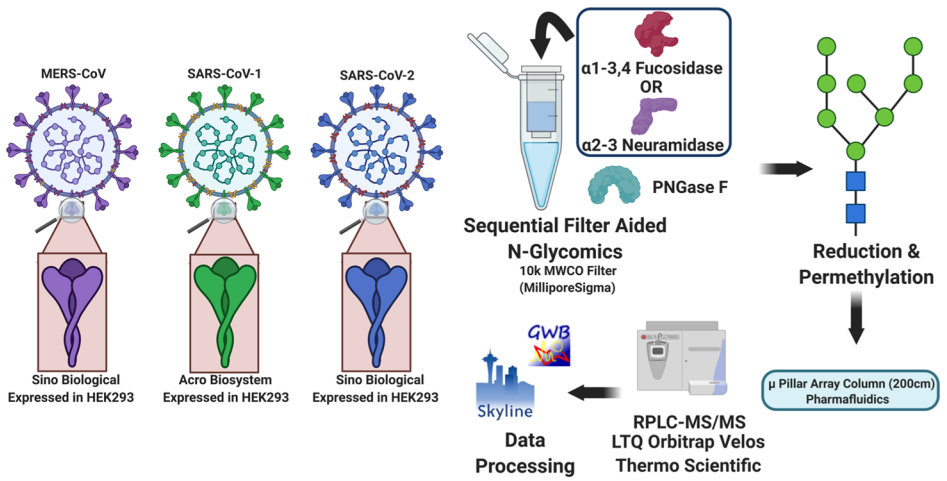


Figure 1. Illustration of the glycomic workflow used in this study.

LC–MS/MS Analysis

For the sample validation proteomic experiment, an Ultimate 3000 nano-LC system (Thermo Scientific) coupled to a Q-Exactive HF mass spectrometer (Thermo Scientific) was used. Peptides digested from 2 μg of protein samples were injected. The peptide samples were trapped using a C18 trap (Acclaim PepMap 100, 75 μm × 2 mm, 3 μm, 100 Å, Thermo Scientific), then separated using a C18 capillary column (Acclaim PepMap 100, 75 μm × 500 mm, 2 μm, 100 Å, Thermo Scientific). The separation was achieved at 0.3 μL/min flow rate using mobile-phase solvents A (2% ACN, 98% water with 0.1% FA) and B (100% ACN with 0.1% FA). The gradient was from 0 to 10 min, 2% B; 10 to 150 min, 2−35% B; 150 to 171 min, 35−80% B; 171 to 185 min, 80% B; and 185 to 186 min, 80−2% B. Column wash with 90% B was applied between two sample runs. The full MS resolution was 120,000 with a scan range of 500−2000 m/z. MS^2 resolution was 30,000 with a maximum IT of 50 ms and data-dependent acquisition (DDA) of top 20 most intense ions in the full MS scan. Stepped collision energies of 20, 30, and 40 were applied for HCD fragmentation.

To perform the isomeric separation of *N*-glycans derived from S proteins that originated from three different coronaviruses, an Ultimate 3000 nano-LC system (Thermo Scientific) coupled to a LTQ Velos Orbitrap mass spectrometer (Thermo Scientific) was used. The separation column was a 200 cm Micro Pillar Array Column (200 cm μPAC, PharmaFluidics). 1 μg of the sample amount derived from the starting material was injected into the system. The samples were online-purified using a μPAC trapping column by loading the solvent (98% water, 2% ACN, and 0.1% formic acid) for 4 min at 10 μL/min and then subjected to a μPAC column for separation. The separation of the samples was achieved at 300 nL/min flow rate, using the following gradient: mobile phase A was 98% water, 2% ACN, and 0.1% formic acid, while mobile phase B was 80% ACN in aqueous solution with 0.1% formic acid. 40% B from 0 to 4 min; 40 to 64% B from 4 to 244 min; 64 to 97.5% B from 244 to 249 min; 97.5% B from 249 to 283 min; 97.5 to 40% B from 283 to 288 min; and 40% B from 288 to 322 min.

The outlet of the LC column was interfaced with a mass spectrometer through a nanoESI source with 1.6 kV of electrospray ionization (ESI) voltage. A mass resolution of 60,000 was employed for full MS with a scan range of 700−2000 m/z. DDA was employed for the collisional induced dissociation (CID). The top eight most intense precursor ions were selected from each full scan for the subsequent MS2 analysis. The CID-normalized collision energy was set to 35% with 3 m/z isolation width and an activation Q value of 0.25. Dynamic exclusion was enabled with a repeat count of 2, a repeat duration of 30 s, and an exclusion duration of 30 s.

Data Processing

For the proteomic experiment, data collected were subject to MaxQuant for protein identification and quantitation. Proteins were searched against the SARS-CoV-1 S protein, SARS-CoV-2 S protein, MERS-CoV S protein, and human proteome UP000005640. The identification FDR was 0.01. The MaxQuant score was set to 60. The assignment of glycoproteins was based on UniprotKB. The abundance of the viral spike proteins relative to human proteins was calculated. Human glycoproteins which might affect the quantitation of viral spike protein *N*-glycans were investigated.

For the glycomic experiment, acquired.raw files were processed *via* Skyline (MacCoss Lab Software) with an in-house glycan library. All possible m/z values of each glycan adduct forms were evaluated manually. The relative quantitation of glycans was performed using Microsoft Excel. Chromatograms and MS^2 spectra were generated with Skyline and XCalibur Qual Browser 4.3 (Thermo Scientific). MS^2 spectra were carefully examined with GlycoWorkbench 2. Other figures were made with GraphPad Prism 8.4.3. A workflow scheme was made with biorender.com.

DATA AVAILABILITY

The raw data are available on GlycoPOST under announced ID GPST000204. The preview can be accessed *via* the link: https://glycopost.glycosmos.org/preview/195490910360e39080eada8. The pin code is 7340.

RESULTS AND DISCUSSION

S1 Protein Total Glycosylation Comparison between the Three Coronaviruses

In this work, *N*-glycans derived from S1 protein of three different coronaviruses, MERS-CoV, SARS-CoV-1, and SARS-CoV-2, that were expressed in HEK293 cells were directly compared. It is noteworthy to discuss that because they were expressed in the same cell line, the only difference between the

https://doi.org/10.1021/acs.jproteome.1c00323
J. Proteome Res. XXXX, XXX, XXX−XXX

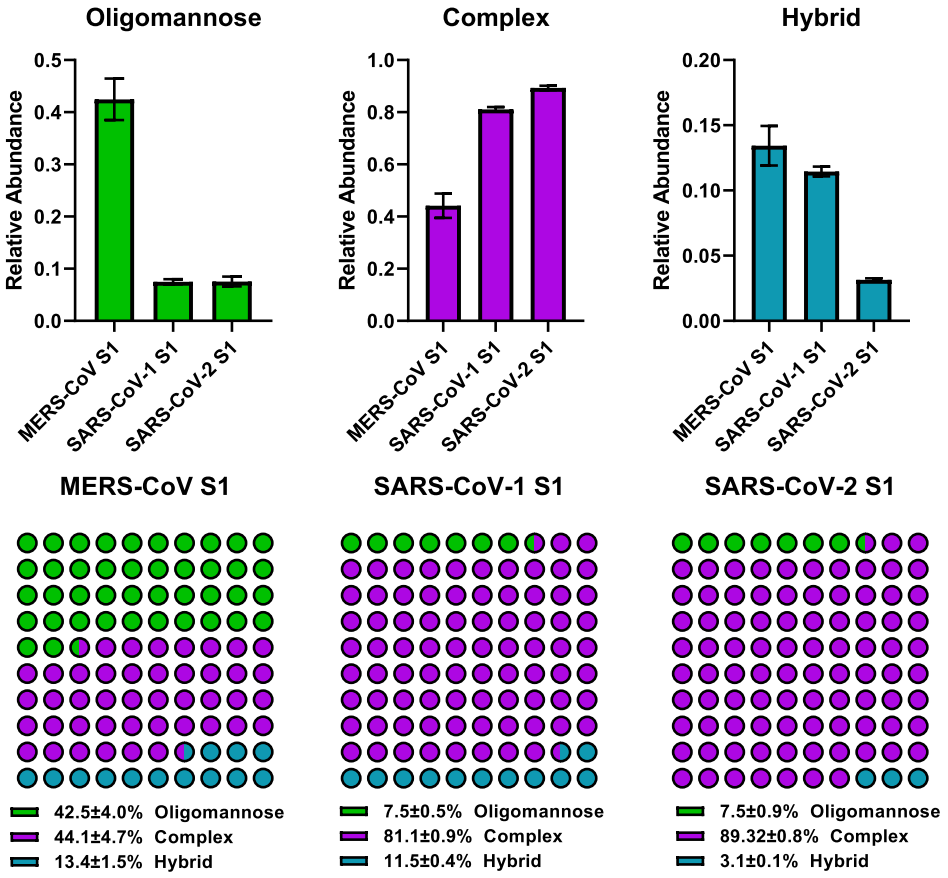


Figure 2. Bar graph depicting the distribution of oligomannose, complex, and hybrid type *N*-glycans derived from S1 proteins among MERS-CoV, SARS-CoV-1, and SARS-CoV-2. Error bars represent standard deviation (n = 3).

proteins is the amino acid sequence (Figure S1) and glycosylation sites. Therefore, the difference in glycan expressions and/or isomer expressions is not a bias caused by the cells. In terms of the sample preparation, although these proteins were purchased from the vendors, the purity of the protein was not assumed, and a bottom-up proteomic experiment was performed to verify the protein in each sample, as shown in Figure S2. The proteomic experiment results showed that the majority of the proteins in each sample belong to each coronavirus S1 protein (>94.7%). To ensure that there are no other glycans influencing the glycomic experiment, the proteomic experimental result was filtered for glycoproteins, in which the result displayed that >98.6% of the sample were in fact S1 proteins from coronaviruses. This demonstrates that the glycans derived from S1 proteins from each coronavirus mostly originated from S1 proteins and not from interferent proteins.

The glycomic experiment was performed with sequential filter-aided *N*-glycomics (FANGs) to simultaneously purify the glycoproteins, apply appropriate exoglycosidases (α1-3,4 fucosidase and α2-3 neuraminidase S), and release *N*-glycans. Permethylation was utilized to enhance ionization as well as enable RPLC separation using a 200 cm micropillar array column (μPAC, Pharmafluidics) coupled with a high-resolution mass spectrometer. This is the first time a μPAC column was used to analyze glycans. The experimental scheme is depicted in Figure 1.

First, the types of *N*-glycans derived from S1 proteins of the three different coronaviruses are examined, oligomannose, complex, and hybrid, as depicted in Figure 2. Oligomannose glycans are composed of Man5, Man6, Man7, Man8, and Man9. MERS-CoV S1 has shown high abundance of oligomannose, while SARS-CoV-1 S1 and SARS-CoV-2 S1 showed a relatively similar expression of oligomannose abundance. However, the complex-type glycans, which contain sialic acid and fucose residues without any extended mannoses, have shown in high abundance in both SARS S1 glycoproteins. SARS-CoV-2 S1 protein, in particular, showed a significantly higher amount of complex glycans while exhibiting a lower abundance for hybrid-type glycans, which are the fusion of oligomannose and complex-type glycans. Based on this, one can say that both SARS S1 proteins shared a comparable amount of oligomannose glycans (SARS-CoV-1, 7.48% *vs* SARS-CoV-2, 7.54%), although there was a significant difference in the abundance of hybrid-type glycans (MERS-CoV, 13.43% *vs* SARS-CoV-1, 11.45% *vs* SARS-CoV-2, 3.14%). At the same time, the MERS-CoV S1 protein exhibited a significantly lower amount of complex glycan type (MERS-CoV, 44.11% *vs* SARS-CoV-1, 81.08% *vs* SARS-CoV-2, 89.32%). However, this result may appear contradictory to other recent works by Watanabe *et al.*[26] and Zhao *et al.*[14] This incongruity is anticipated due to differences in the scope of the studies. The current work investigates the isomeric glycosylation of the S1 subunit of the S protein, while the other two studies explore glycosylation in both S1 and S2 subunits. Thus, making any direct comparisons between this study and the other two studies might be inappropriate. Moreover, the immune response against the coronavirus generates most of the neutralizing antibodies against RBD, which is in the S1 subunit. Thus, S1 might be a better

https://doi.org/10.1021/acs.jproteome.1c00323
J. Proteome Res. XXXX, XXX, XXX−XXX

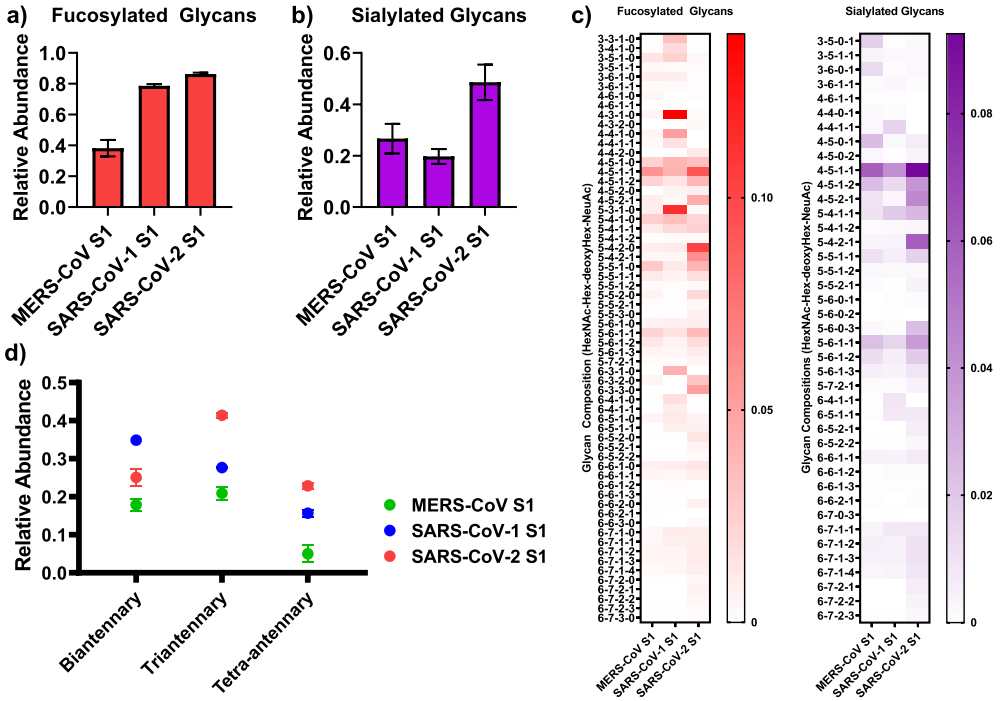


Figure 3. Relative abundance of (a) fucosylated glycans and (b) sialylated glycans derived from S1 proteins among MERS-CoV, SARS-CoV-1, and SARS-CoV-2. (c) Heatmaps of individual glycans from S1 proteins illustrating relative abundance. (d) Comparison of bi-, tri-, and tetra-antennary glycan abundance from S1 proteins among MERS-CoV, SARS-CoV-1, and SARS-CoV-2. Error bars represent standard deviation (n = 3).

investigation model for designing therapeutics.[9] The S1 protein is known to be the initial contact point for viral interaction with the host, while the S2 protein is involved in the viral membrane fusion process,[27] which is why the S1 protein was chosen for glycosylation analysis. It is also noteworthy to mention that glycan analysis by Watanabe *et al.*[26] was performed through the fluorescence labeling of glycans and separation was done by hydrophilic interaction liquid chromatography coupled with a fluorescence detector, whereas this work and Zhao *et al.*[14] performed permethylation of glycans and separation was done with RPLC coupled with a high-resolution mass spectrometer.

Fucose and sialic acid containing *N*-glycans are also examined, and in Figure 3a, b, a comparison of the total relative abundance of fucosylated and sialylated glycan among the three coronavirus S1 proteins is shown. Notably, both SARS-CoV S1 proteins showed a high abundance of fucosylated glycans. *N*-Glycans derived from SARS-CoV-2 S1, in particular, were mostly fucosylated with a relative abundance of 0.86 and in comparison, 0.79 of *N*-glycans derived from SARS-CoV-1 S1 were fucosylated. More significantly, only 0.38 of *N*-glycans from MERS-CoV S1 were fucosylated. This is due to the high abundance of the oligomannose type of glycans shown in MERS-CoV S1. A high amount of oligomannose also affected the abundance of *N*-glycans with sialic acids. MERS-CoV S1 and SARS-CoV-1 S1 showed 0.27 and 0.20 relative abundances, respectively. SARS-CoV-2 S1, however, revealed nearly half of all glycans (0.49) with sialic acid, which is significantly higher than both MERS-CoV S1 and SARS-CoV-1 S1. This analysis suggests significantly more complex glycans being expressed on SARS-CoV-2 S1 compared to the other two coronaviruses' S1 protein. Individual fucosylated and sialylated glycans were also assessed, which is shown in Figure 3c in the form of a heatmap with the glycan composition listed on the y-axis and different coronavirus S1 proteins shown on the x-axis to evaluate the similarity and disparity of the glycans being expressed on each protein. For example, HexNAc4Hex5deoxyHex1 could be a comparable expression, but HexNAc5Hex4deoxyHex2 was significantly more expressed in SARS-CoV-2 S1 than the others. As for the sialylated glycans, HexNAc5Hex6deoxyHex1NeuAc1 showed a similar expression across the three proteins, but HexNAc4Hex5deoxyHex1NeuAc1 was much higher in abundance for SARS-CoV-2 S1. The comparison of individual fucosylated and sialylated glycans indicates the disparity of them being expressed across the three coronavirus S1 proteins and highlights that many common glycans are being shared between the three S1 proteins. Glycans with different numbers of antennas are also assessed, as depicted in Figure 3d. For this analysis, glycans with four HexNAc were accounted as biantennary, five HexNAc as triantennary, and six HexNAc as tetra-antennary. The MERS-CoV S1 protein displayed a low overall relative abundance due to the high expression of the oligomannose type and the lack of complex-type glycans, as mentioned above. Biantennary *N*-glycans were the most abundant in the SARS-CoV-1 S1 protein, followed by triantennary, and then tetra-antennary. The SARS-CoV-2 S1 protein, however, had triantennary glycans (0.41) the most, followed by approximately the same amount of biantennary (0.25) and tetra-antennary (0.23) glycans. This also alludes to the fact that glycosylation in the S1 protein from SARS-CoV-2 might be more complex than its counterparts; SARS-CoV-1 and MERS-CoV.

N-Glycan Isomers Derived from S1 Proteins from Three Coronaviruses

Glycan isomers are an essential feature of glycosylation, especially contributing to the virion and host interaction. One example of glycan isomers being associated with the viral interaction of the host is the influenza virus. Sialic acid linkage

https://doi.org/10.1021/acs.jproteome.1c00323
J. Proteome Res. XXXX, XXX, XXX−XXX

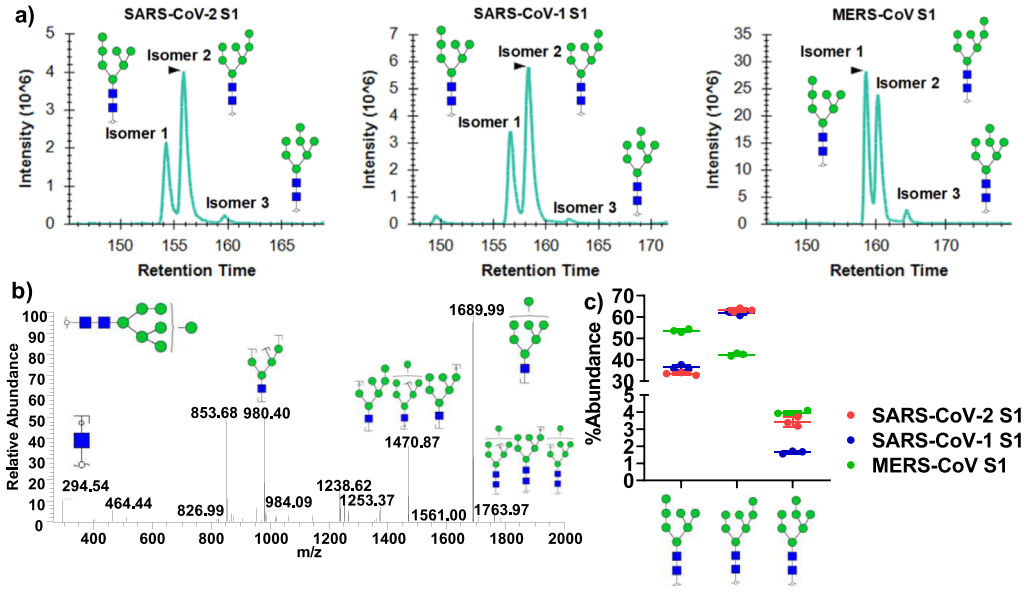


Figure 4. (a) EICs of Man7 isomers derived from S1 proteins among MERS-CoV, SARS-CoV-1, and SARS-CoV-2. (b) MS^2 spectrum of Man7, evidence of glycan identity. (c) Man7 glycan isomer distribution among MERS-CoV, SARS-CoV-1, and SARS-CoV-2 S1 proteins.

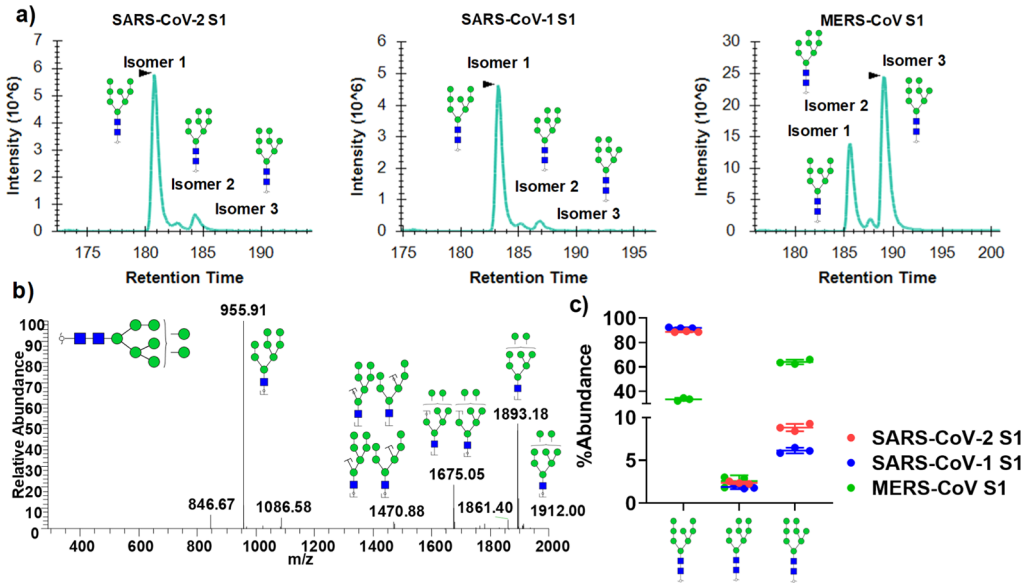


Figure 5. (a) EICs of Man8 isomers derived from S1 proteins among MERS-CoV, SARS-CoV-1, and SARS-CoV-2. (b) MS^2 spectrum of Man8, evidence of glycan identity. (c) Man8 glycan isomer distribution among MERS-CoV, SARS-CoV-1, and SARS-CoV-2 S1 proteins.

isomers have been correlated with effect of the influenza virus infection that varies depending on the host.[18] Based on this information, we hypothesized that the high communicability of the SARS-CoV-2 virus might be associated with aberrant glycan isomers being expressed on the S1 protein. To assess this postulation, we employed a 200 cm μPAC column from Pharmafluidics in anticipation that the length of the column could perhaps be able to resolve permethylated glycan isomers. Figure 4a depicts the separation of Man7 isomers on 200 cm μPAC with MS^2 spectra (Figure 4b) to confirm the structure and isomer distribution (Figure 4c) and is shown to evaluate the isomeric glycan expression differences among the S1 proteins. Isomer identification was deduced by a previous work which suggested that Man7 isomer separation is the reverse order of elution of separation on a porous graphitized carbon column.[28] There was an apparent disparity of the Man7 isomer distribution between the SARS S1 proteins and MERS-CoV S1, showing a significant difference in isomer 1 and 2 glycan distributions. Figure 5a illustrates the separation of Man8 isomers. Interestingly, the distribution of Man8 isomers (Figure 5c) from both SARS S1 proteins were comparable, while the isomers of the MERS-CoV S1 protein were remarkably distinct, same as Man7 isomers.

Fucosylated and sialylated glycan isomers were also investigated among the three S1 proteins. Few of the most abundant glycans were assessed for the presence of isomers. Extracted ion chromatograms (EIC) of biantennary mono-fucosylated and monosialylated glycan isomers are shown in Figure 6a. As depicted, both sialic acid linkage isomers and fucose positional isomers were separated. This was confirmed using exoglycosidases, α2-3 neuraminidase and α1-3,4 fucosidase, as shown in Figure 6b. EIC of the α2-3

https://doi.org/10.1021/acs.jproteome.1c00323
J. Proteome Res. XXXX, XXX, XXX−XXX

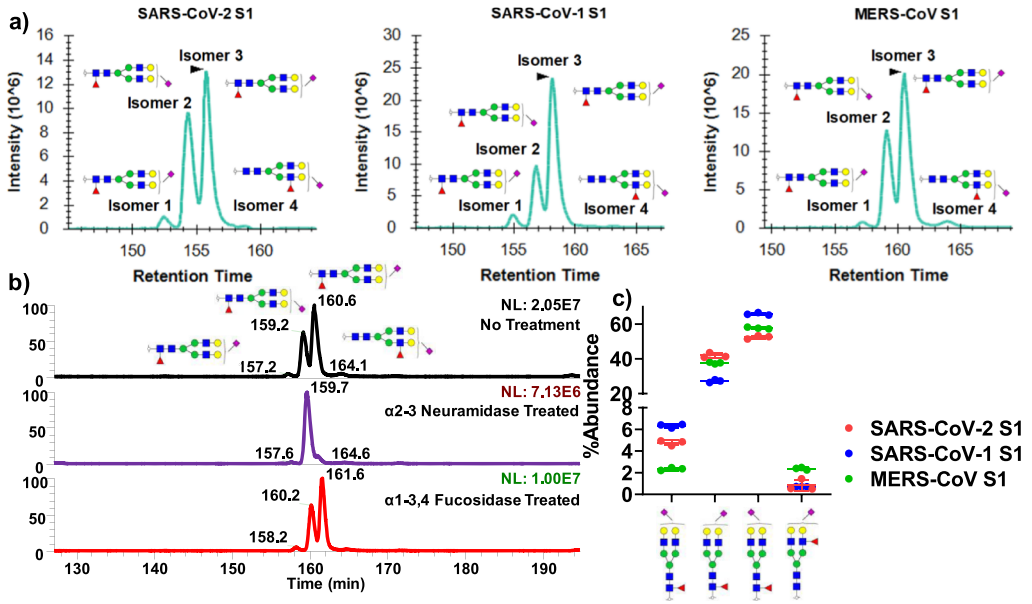


Figure 6. (a) EICs of biantennary monofucosylated monosialylated glycan isomers derived from S1 proteins among MERS-CoV, SARS-CoV-1, and SARS-CoV-2. (b) EICs of biantennary monofucosylated monosialylated glycan isomers with α2-3 neuraminidase and α1-3,4 fucosidase. Disappearance of the peaks indicates the cleavage of specific linkage isomers. (c) Biantennary monofucosylated monosialylated glycan isomer distribution among MERS-CoV, SARS-CoV-1, and SARS-CoV-2 S1 proteins.

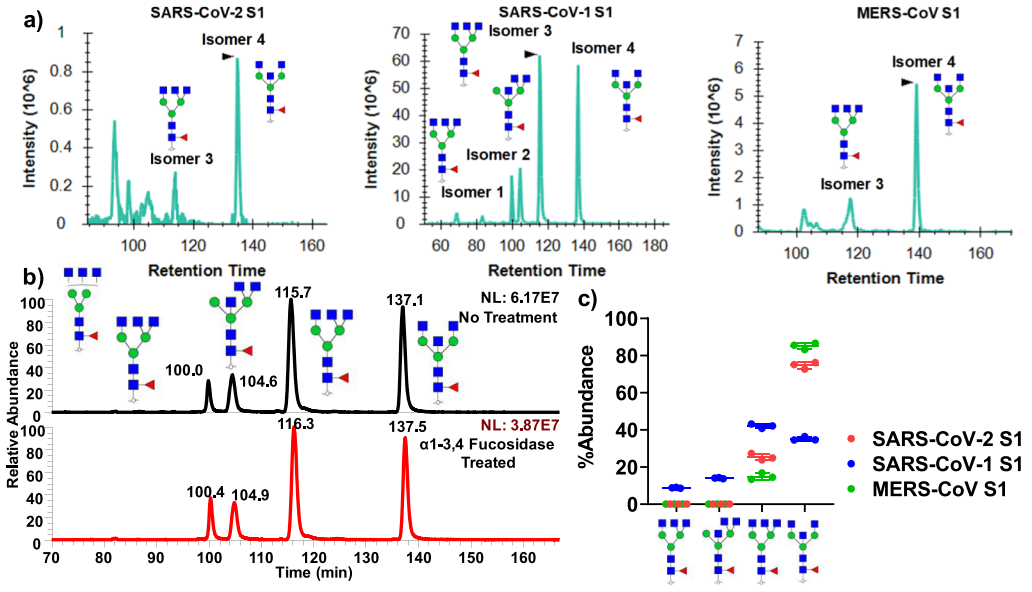


Figure 7. (a) EICs of GlcNAc5Man3deoxyHex1 glycan isomers derived from S1 proteins among MERS-CoV, SARS-CoV-1, and SARS-CoV-2. (b) EICs of GlcNAc5Man3deoxyHex1 glycan isomers with α2-3 neuraminidase and α1-3,4 fucosidase. There was a lack of isomers 1 and isomer 2 presented in SARS-CoV-2 and MERS-CoV S1 proteins. (c) GlcNAc5Man3deoxyHex1 glycan isomer distribution among MERS-CoV, SARS-CoV-1, and SARS-CoV-2 S1 proteins. Isomers are assigned based on the MS^2 spectra shown in Figure S6 and S7.

neuraminidase-treated sample only displays α2-6 sialic acid-linked glycans. Because the isomer 2 peak did not disappear, it was determined that isomer 2 had α2-6 sialic acid linkage. Isomers 1 and 3, therefore, were α2-3 sialic acid linkages; however, according to the EIC from the α1-3,4 fucosidase-treated sample, they were both α1-6 core fucose, which suggests that isomers 1 and 3 could be different arm linkages (α-3 or α-6). Isomer 4 was determined as branch fucose because the isomer 4 peak faded when it was treated with α1-3,4 fucosidase. Biantennary monofucosylated and disialylated glycan are also assessed using the same technique illustrated in Figure S3. Two chromatographic peaks are observed in Figure S3a; however, it is not feasible to determine the exact structure of the isomers possibly due to the lack of separation or lack of either two α2-3 or two α2-6 sialic acid linkage isomers. Neither peak from the α1-3,4 fucosidase-treated chromatogram diminished, which suggests that both peaks were α1-6 core fucose. Peaks from the α2-3 neuraminidase-treated sample also did not fade; however, a significant reduction in intensity insinuates that both peaks contain α2-3-linked sialic acids. Although it could be an indication that the sialic acid linkage isomers were not fully resolved or there was a lack of isomers, SARS-CoV-2 S1 and MERS-CoV S1 are comparable in distribution, while SARS-CoV-1 S1 showed a little discrepancy

https://doi.org/10.1021/acs.jproteome.1c00323
J. Proteome Res. XXXX, XXX, XXX−XXX

(Figure S3b). More complex glycans, such as tetra-antennary monofucosylated and tetrasialylated glycan isomers, are evaluated, as shown in Figure S4, where four isomers were illustrated. The representation of α1-3,4 fucosidase- and α2-3 neuraminidase-treated samples (Figure S4b), suggesting that all four peaks are characterized as core fucosylated glycans as well as an incomplete resolution among the sialic acid linkage isomers. Notably, isomers 3 and 4 were not determined in the SARS-CoV-1 S1 protein, while their counterparts did, which show similarity in the glycan isomer expression between SARS-CoV-2 S1 and MERS-CoV S1 proteins, as shown in Figure S4c.

Bisecting glycan isomers are also investigated in this study, as shown in Figure 7. EICs of GlcNAc5Man3deoxyHex1 depicted in Figure 7a demonstrated four isomers with SARS-CoV-1 S1, while the other two only exhibited two isomers. α1-3,4 fucosidase was again utilized to determine whether these isomers are branch- or core-fucosylated, as shown in Figure S5. However, the exoglycosidase treatment had no effect on chromatography; thus, these isomers were not branch-fucosylated but core-fucosylated. To determine the GlcNAc positions, MS^2 spectra were utilized (Figure S6). With the anticipation that many fragment ions could be overlapping among the isomers, we employed Glycoworkbench to generate theoretical m/z values of all possible fragments for each isomer. Then, cross-referenced them with the fragment ion list yielded from Xcalibur software to determine the possible diagnostic ions to distinguish nonbisecting and bisecting glycans with stringent parameters with the actual fragment m/z matching the theoretical one within 0.25 Da and a minimum relative abundance of 1%. As depicted in Figure S6, isomers 2 and 4 contain fragment ions matched with bisecting glycans, which meant isomers 1 and 3 are nonbisecting. Then, to assess whether isomers 2 and 4 were of hybrid or complex type, the same procedure was repeated (Figure S7). As a result, isomer 2 was suggested to be the hybrid type where one of the mannoses did not have an extending GlcNAc branch, and isomer 4 was complex where all mannoses had a GlcNAc branch. Furthermore, we have identified other abundant glycan isomers, as shown in Figures S8–S15, among the three S1 proteins.

CONCLUSIONS

In this work, we demonstrated the disparity in glycome from the S1 proteins that originated from three different coronaviruses as well as glycan isomers. The evidence from this work suggests that glycosylation in the SARS-CoV-2 S1 glycoprotein is more complex than SARS-CoV-1 S1 or MERS-CoV S1 glycoproteins. Moreover, analogous and distinctive isomeric glycan distributions among the three glycoproteins were demonstrated. It is worth noting that this study utilized recombinant proteins, which are artificially generated and purified in a lab. Most of the work regarding SARS-CoV-2 was performed with recombinant proteins; however, a recent work by Yao *et al.*[29] characterized the SARS-CoV-2 virus as well as the spike proteins using the native virus, which was isolated from a patient. In comparison with the work by Watanabe *et al.*,[26] the authors claimed some resemblances in glycan compositions between the native virion spike protein and the recombinant spike protein, while native virion spike proteins are surrounded by much more "bulkier" glycans and more complex glycans. More recently, Brun *et al.* showed the differences in the site-specific glycosylation of spike proteins between the recombinant, virus-derived, and vaccine antigens.[30] These pieces of evidence present the demand for isomeric glycan analysis of native virion S proteins for better vaccine development and perhaps a better rapid testing strategy due to the roles that glycan isomers may play in causing infection to the host cells as well as the specificity of antibody- or antigen-based tests. Furthermore, glycan isomers of spike proteins involving the virion mutation should also be considered for the mutated SARS-CoV-2 virus.[31]

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jproteome.1c00323.

Amino acid sequences of the S1 glycoprotein of MERS-CoV, SARS-CoV-1, and SARS-CoV-2 expressed in HEK293 cells; purity of the acquired S1 proteins by MaxQuant; EICs of bi- and tetra-antennary monofucosylated disialylated glycan isomers derived from S1 proteins among MERS-CoV, SARS-CoV-1, and SARS-CoV-2, of glycan isomers with α2-3 neuraminidase and α1-3,4 fucosidase, and the glycan isomer distribution among MERS-CoV, SARS-CoV-1, and SARS-CoV-2 S1 proteins; EICs and tandem MS spectra of GlcNAc5-Man3deoxyHex1 glycan isomers distinguishing hybrid bisecting and complex bisecting *N*-glycan, of Man6, and of GlcNAc5Man3Gal2; and EICs of GlcNAc5Man3-deoxyHex1 glycan isomers, GlcNAc5Man3Gal1deoxy-Hex2, GlcNAc5Man3Gal1deoxyHex2NeuAc1, GlcNAc6Man3Gal1deoxyHex3, GlcNAc4Man3Gal2-deoxyHex2NeuAc1, GlcNAc4Man3deoxyHex1, GlcNAc4Man3Gal1deoxyHex1, GlcNAc6Man3deoxy-Hex1 and Man6 with and without exoglycosidase treatment (PDF)

AUTHOR INFORMATION

Corresponding Author

Yehia Mechref − *Department of Chemistry and Biochemistry, Texas Tech University, Lubbock, Texas 79409-1061, United States;* orcid.org/0000-0002-6661-6073; Phone: 806-742-3059; Email: yehia.mechref@ttu.edu; Fax: 806-742-1289

Authors

Byeong Gwan Cho − *Department of Chemistry and Biochemistry, Texas Tech University, Lubbock, Texas 79409-1061, United States*

Sakshi Gautam − *Department of Chemistry and Biochemistry, Texas Tech University, Lubbock, Texas 79409-1061, United States*

Wenjing Peng − *Department of Chemistry and Biochemistry, Texas Tech University, Lubbock, Texas 79409-1061, United States*

Yifan Huang − *Department of Chemistry and Biochemistry, Texas Tech University, Lubbock, Texas 79409-1061, United States*

Mona Goli − *Department of Chemistry and Biochemistry, Texas Tech University, Lubbock, Texas 79409-1061, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jproteome.1c00323

https://doi.org/10.1021/acs.jproteome.1c00323
J. Proteome Res. XXXX, XXX, XXX−XXX

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by grants from the National Institutes of Health, NIH (1R01GM112490, 1R01GM130091, and 1U01CA225753).

■ REFERENCES

(1) Rothan, H. A.; Byrareddy, S. N. The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. *J. Autoimmun.* **2020**, *109*, 102433.
(2) Iacobucci, G. Covid-19: Lack of capacity led to halting of community testing in March, admits deputy chief medical officer. *Br. Med. J.* **2020**, *369*, m1845.
(3) West, C. P.; Montori, V. M.; Sampathkumar, P. COVID-19 Testing: The Threat of False-Negative Results. *Mayo Clin. Proc.* **2020**, *95*, 1127−1129.
(4) Kim, Y. C.; Dema, B.; Reyes-Sandoval, A. COVID-19 vaccines: breaking record times to first-in-human trials. *npj Vaccines* **2020**, *5*, 34.
(5) Shin, M. D.; Shukla, S.; Chung, Y. H.; Beiss, V.; Chan, S. K.; Ortega-Rivera, O. A.; Wirth, D. M.; Chen, A.; Sack, M.; Pokorski, J. K.; Nicole, F. S. COVID-19 vaccine development and a potential nanomaterial path forward. *Nat. Nanotechnol.* **2020**, *15*, 646−655.
(6) Chauhan, G.; Madou, M. J.; Kalra, S.; Chopra, V.; Ghosh, D.; Martinez-Chapa, S. O. Nanotechnology for COVID-19: therapeutics and vaccine research. *ACS Nano* **2020**, *14*, 7760−7782.
(7) Cui, J.; Li, F.; Shi, Z.-L. Origin and evolution of pathogenic coronaviruses. *Nat. Rev. Microbiol.* **2019**, *17*, 181−192.
(8) De Wit, E.; Van Doremalen, N.; Falzarano, D.; Munster, V. J. SARS and MERS: recent insights into emerging coronaviruses. *Nat. Rev. Microbiol.* **2016**, *14*, 523.
(9) Barnes, C. O.; West, A., Jr; Huey-Tubman, K.; Hoffmann, M.; Sharaf, N.; Hoffman, P.; Koranda, N.; Gristick, H.; Gaebler, C.; Muecksch, F. Structures of human antibodies bound to SARS-CoV-2 spike reveal common epitopes and recurrent features of antibodies. *Cell* **2020**, *182*, 828−842.
(10) Feng, W.; Newbigging, A. M.; Le, C.; Pang, B.; Peng, H.; Cao, Y.; Wu, J.; Abbas, G.; Song, J.; Wang, D.-B. Molecular Diagnosis of COVID-19: Challenges and Research Needs. *Anal. Chem.* **2020**, *92*, 10196−10209.
(11) 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. Research and development on therapeutic agents and vaccines for COVID-19 and related human coronavirus diseases. *ACS Cent. Sci.* **2020**, *6*, 315−331.
(12) 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**, *30*, 981−988.
(13) Watanabe, Y.; Allen, J. D.; Wrapp, D.; McLellan, J. S.; Crispin, M. Site-specific glycan analysis of the SARS-CoV-2 spike. *Science* **2020**, *369*, 330−333.
(14) Zhao, P.; Praissman, J. L.; Grant, O. C.; Cai, Y.; Xiao, T.; Rosenbalm, K. E.; Aoki, K.; Kellman, B. P.; Bridger, R.; Barouch, D. H.; Brindley, M. A.; Lewis, N. E.; Tiemeyer, M.; Chen, B.; Woods, R. J.; Wells, L. Virus-Receptor Interactions of Glycosylated SARS-CoV-2 Spike and Human ACE2 Receptor. *Cell Host Microbe* **2020**, *28*, 586−601.
(15) Zhou, D.; Tian, X.; Qi, R.; Peng, C.; Zhang, W. Identification of 22 N-glycosites on spike glycoprotein of SARS-CoV-2 and accessible surface glycopeptide motifs: implications for vaccination and antibody therapeutics. *Glycobiology* **2020**, *31*, 69−80.
(16) Antonopoulos, A.; Broome, S.; Sharov, V.; Ziegenfuss, C.; Easton, R. L.; Panico, M.; Dell, A.; Morris, H. R.; Haslam, S. M. Site-specific characterisation of SARS-CoV-2 spike glycoprotein receptor binding domain. *Glycobiology* **2020**, *31*, 181−187.
(17) Broszeit, F.; Tzarum, N.; Zhu, X.; Nemanichvili, N.; Eggink, D.; Leenders, T.; Li, Z.; Liu, L.; Wolfert, M. A.; Papanikolaou, A. N-glycolylneuraminic acid as a receptor for influenza A viruses. *Cell Rep.* **2019**, *27*, 3284−3294.e6.
(18) Gagneux, P.; Cheriyan, M.; Hurtado-Ziola, N.; van der Linden, E. C. B.; Anderson, D.; McClure, H.; Varki, A.; Varki, N. M. Human-specific regulation of α2−6-linked sialic acids. *J. Biol. Chem.* **2003**, *278*, 48245−48250.
(19) Tortorici, M. A.; Walls, A. C.; Lang, Y.; Wang, C.; Li, Z.; Koerhuis, D.; Boons, G.-J.; Bosch, B.-J.; Rey, F. A.; de Groot, R. J. Structural basis for human coronavirus attachment to sialic acid receptors. *Nat. Struct. Mol. Biol.* **2019**, *26*, 481−489.
(20) Zhou, S.; Veillon, L.; Dong, X.; Huang, Y.; Mechref, Y. Direct comparison of derivatization strategies for LC-MS/MS analysis of N-glycans. *Analyst* **2017**, *142*, 4446−4455.
(21) de Haan, N.; Yang, S.; Cipollo, J.; Wuhrer, M. Glycomics studies using sialic acid derivatization and mass spectrometry. *Nat. Rev. Chem.* **2020**, *4*, 229−242.
(22) Zhou, S.; Wooding, K. M.; Mechref, Y. Analysis of permethylated glycan by liquid chromatography (LC) and mass spectrometry (MS), *High-Throughput Glycomics and Glycoproteomics*; Springer, 2017, pp 83−96.
(23) Dong, X.; Huang, Y.; Cho, B. G.; Zhong, J.; Gautam, S.; Peng, W.; Williamson, S. D.; Banazadeh, A.; Torres-Ulloa, K. Y.; Mechref, Y. Advances in mass spectrometry-based glycomics. *Electrophoresis* **2018**, *39*, 3063−3081.
(24) Gautam, S.; Peng, W.; Cho, B.; Huang, Y.; Banazadaeh, A.; Yu, A.; Dong, X.; Mechref, Y. Glucose Unit Index (GUI) of Permethylated Glycans For Effective Identification of Glycans and Glycan Isomers. *Analyst* **2020**, *145*, 6656−6667.
(25) Kang, P.; Mechref, Y.; Novotny, M. V. High-throughput solid-phase permethylation of glycans prior to mass spectrometry. *Rapid Commun. Mass Spectrom.* **2008**, *22*, 721−734.
(26) Watanabe, Y.; Berndsen, Z. T.; Raghwani, J.; Seabright, G. E.; Allen, J. D.; Pybus, O. G.; McLellan, J. S.; Wilson, I. A.; Bowden, T. A.; Ward, A. B.; Crispin, M. Vulnerabilities in coronavirus glycan shields despite extensive glycosylation. *Nat. Commun.* **2020**, *11*, 2688.
(27) 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**, *181*, 281−292.
(28) Zhou, S.; Huang, Y.; Dong, X.; Peng, W.; Veillon, L.; Kitagawa, D. A. S.; Aquino, A. J. A.; Mechref, Y. Isomeric Separation of Permethylated Glycans by Porous Graphitic Carbon (PGC)-LC-MS/MS at High Temperatures. *Anal. Chem.* **2017**, *89*, 6590−6597.
(29) Yao, H.; Song, Y.; Chen, Y.; Wu, N.; Xu, J.; Sun, C.; Zhang, J.; Weng, T.; Zhang, Z.; Wu, Z.; Cheng, L.; Shi, D.; Lu, X.; Lei, J.; Crispin, M.; Shi, Y.; Li, L.; Li, S. Molecular architecture of the SARS-CoV-2 virus. *Cell* **2020**, *183*, 730−738.
(30) Brun, J.; Vasiljevic, S.; Gangadharan, B.; Hensen, M.; Chandran, V. A.; Hill, M. L.; Kiappes, J. L.; Dwek, R. A.; Alonzi, D. S.; Struwe, W. B.; Zitzmann, N. Assessing Antigen Structural Integrity through Glycosylation Analysis of the SARS-CoV-2 Viral Spike. *ACS Cent. Sci.* **2021**, *7*, 586−893.
(31) Korber, B.; Fischer, W. M.; Gnanakaran, S.; Yoon, H.; Theiler, J.; Abfalterer, W.; Hengartner, N.; Giorgi, E. E.; Bhattacharya, T.; Foley, B.; Hastie, K. M.; Parker, M. D.; Partridge, D. G.; Evans, C. M.; Freeman, T. M.; de Silva, T. I.; Sheffield, C.-G. G.; McDanal, C.; Perez, L. G.; Tang, H.; Moon-Walker, A.; Whelan, S. P.; LaBranche, C. C.; Saphire, E. O.; Montefiori, D. C. Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. *Cell* **2020**, *182*, 812−827 e19.

https://doi.org/10.1021/acs.jproteome.1c00323
J. Proteome Res. XXXX, XXX, XXX−XXX