O-Glycosylation Landscapes of SARS-CoV-2 Spike Proteins

Yong Zhang1*, Wanjun Zhao2, Yonghong Mao2, Yaohui Chen3, Shanshan Zheng1, Wei Cao1, Jingqiang Zhu2, Liqiang Hu1, Meng Gong1, Jingqiu Cheng1* and Hao Yang1*

1Key Laboratory of Transplant Engineering and Immunology, MOH, Frontiers Science Center for Disease-related Molecular Network, Institutes for Systems Genetics, West China Hospital, Sichuan University, Chengdu, China, 2Department of Thyroid Surgery, West China Hospital, Sichuan University, Chengdu, China, 3Institute of Thoracic Oncology, West China Hospital, Sichuan University, Chengdu, China

The densely glycosylated spike (S) proteins that are highly exposed on the surface of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) facilitate viral attachment, entry, and membrane fusion. We have previously reported all the 22 N-glycosites and site-specific N-glycans in the S protein protomer. Herein, we report the O-glycosylation landscapes of SARS-CoV-2 S proteins, which were characterized through high-resolution mass spectrometry. Following digestion with trypsin and trypsin/Glu-C, and de-N-glycosylation using PNGase F, we determined the GalNAc-type O-glycosylation pattern of S proteins, including O-glycosites and the six most common O-glycans occupying them, via Byonic identification and manual validation. Finally, 255 intact O-glycopeptides composed of 50 peptide sequences and 43 O-glycosites were discovered by higher energy collision-induced dissociation (HCD), and three O-glycosites were confidently identified by electron transfer/higher energy collision-induced dissociation (EThcD) in the insect cell-expressed S protein. Most glycosites were modified by non-sialylated O-glycans such as HexNAc(1) and HexNAc(1)Hex (1). In contrast, in the human cell-expressed S protein S1 subunit, 407 intact O-glycopeptides composed of 34 peptide sequences and 30 O-glycosites were discovered by HCD, and 11 O-glycosites were unambiguously assigned by EThcD. However, the measurement of O-glycosylation occupancy hasn’t been made. Most glycosites were modified by sialylated O-glycans such as HexNAc(1)Hex (1)NeuAc (1) and HexNAc(1)Hex (1)NeuAc (2). Our results reveal that the SARS-CoV-2 S protein is an O-glycoprotein; the O-glycosites and O-glycan compositions vary with the host cell type. These comprehensive O-glycosylation landscapes of the S protein are expected to provide novel insights into the viral binding mechanism and present a strategy for the development of vaccines and targeted drugs.

Keywords: SARS-CoV-2, spike protein, O-glycosylation, mass spectrometry, EThcD fragmentation

INTRODUCTION

The spike (S) protein of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is an extensively N-glycosylated protein (Watanabe et al., 2020) that protrudes from the virus surface and binds to the angiotensin-converting enzyme 2 (ACE2) receptor on host cells to mediate cell entry (Wrapp et al., 2020). All 22 N-glycosites and N-glycans attached to asparagine (Asn, N) in a recombinant S protein protomer expressed in human and insect cells have been identified using
high-resolution liquid chromatography–tandem mass spectrometry (LC-MS/MS) (Lenza et al., 2020; Rosenbalm et al., 2020; Walls et al., 2020; Xu et al., 2020; Yan et al., 2020; Zhang et al., 2020; Wang et al., 2021; Zhou et al., 2021). These N-glycosites are preferentially distributed in two functional subunits responsible for receptor binding (S1 subunit) and membrane fusion (S2 subunit) (Zhang et al., 2020). Site-specific N-glycosylation analysis can provide valuable insights into the infection mechanism and present a strategy for the development of vaccines (Grant et al., 2020).

Unlike N-glycosylation, O-glycosylation is initiated by the α-glycosidic attachment of N-acetylglactosamine (GalNAc) to the hydroxyl group of serine (Ser, S) or threonine (Thr, T), which contains eight types of core structures (Core-1 to Core-8 O-glycans), and is involved in a variety of biological functions, such as the mediation of pathogenic binding to human receptors (Mayr et al., 2018; Shajahan et al., 2020a). Moreover, O-glycosylation can influence proteolysis during antigen processing, which could prevent the formation of glycopeptides for further presentation to major histocompatibility complex (MHC) and the elicitation of immune response (Wolffert and Boons, 2013). The S protein O-glycosites of SARS-CoV-2 have been predicted using computational analysis (Uslupehliyan and Sener, 2020), and Shajahan et al. (2020) identified two O-glycosites (T323 and S325) using LC-MS/MS (Shajahan et al., 2020b). However, O-glycosylation often occurs in a cluster. Hence, we believe that there are many O-glycosites that have not been discovered as deciphering protein O-glycosylation remains a big challenge. The comprehensive O-glycosylation analysis cannot be performed without appropriate sample preprocessing, analysis methods, and software (King et al., 2017; Qin et al., 2017; Yang et al., 2017; Yang et al., 2018; Ye et al., 2019; Park et al., 2020; Dong et al., 2021).

In the present study, we characterized the intact O-glycopeptides of recombinant SARS-CoV-2 S proteins expressed in human and insect cells, using LC-MS/MS. Based on a complementary enzyme digestion strategy, we identified large-scale O-glycosites and their corresponding O-glycans in the recombinant S proteins. The heterogeneity by different glycoforms of S protein S1 subunits expressed in human and insect cells was resolved and compared. Detailed O-glycosylation profiles of S proteins are complementary to the N-glycosylation profiles and may help in the development of vaccines and therapeutic drugs.

**EXPERIMENTAL SECTION**

**Materials and Chemicals**

Dithiothreitol (DTT), iodoacetamide (IAA), formic acid (FA), trifluoroacetic acid (TFA), Tris base, and urea were purchased from Sigma (St. Louis, MO, United States). Acetonitrile (ACN) was purchased from Merck (Darmstadt, Germany). Zwitterionic hydrophilic interaction liquid chromatography (ZIC-HILIC) materials were purchased from Fresh Bioscience (Shanghai, China). The C18 and C8 membrane were purchased from Agela Technologies (Tianjin, China). Recombinant SARS-CoV-2 S protein (S1+S2 ECD, His tag) expressed by insect cells (High Five) via a baculovirus, and S protein (S1, His tag) expressed by human embryonic kidney (HEK293) cells were purchased from Sino Biological (Beijing, China). Codon-optimized DNA sequences encoding the SARS-CoV-2 S protein subunits were cloned into pCMV3-C-His and a baculovirus vector with a poly-histidine tag at the C terminus for recombinant expression of these proteins in human and insect cells, respectively. Sequencing-grade trypsin and Glu-C were obtained from Enzyme & Spectrum (Beijing, China). A quantitative colorimetric peptide assay kit was purchased from Thermo Fisher Scientific (Waltham, MA, United States). Deionized water was prepared using a Milli-Q system (Millipore, Bedford, MA, United States). All other chemicals and reagents of the best available grade were purchased from Sigma-Aldrich or Thermo Fisher Scientific.

**Protein digestion**

Recombinant S proteins were proteolyzed using an insulin protease digestion protocol. In brief, 50 μg of protein were dissolved in 100 μl of 50 mM NH₄HCO₃ buffer (pH = 8.5) and heated to denature for 10 min at 95°C. After reduction by DTT (20 mM) for 45 min at 56°C and alkylation with IAA (50 mM) for 1 h at 25°C in the dark, 2 μg of protease (trypsin or trypsin/Glu-C (w/w = 1:1)) was added to the tube and incubated for 16 h at 37°C. Peptides were loaded in a pipette tip which was packed with a C18 membrane. After washing three times using 70 μl of 2% acetonitrile/98% water/0.1% formic acid. Peptides bound to the C18 membrane were eluted three times with 70 μl of 80% acetonitrile/20% water/0.1% formic acid. The peptide concentration was determined using a peptide assay kit, based on the absorbance measured at 480 nm. The peptide mixtures were freeze-dried for further analysis.

**Enrichment of Intact Glycopeptides and N-Glycan Removal**

Intact N- and O-glycopeptides were enriched with ZIC-HILIC materials. Specifically, 20 μg of peptides was suspended in 100 μl of 80% ACN/0.2% TFA solution. 5 μg of Zic-HILIC was washed three times for 10 min each with 0.1% TFA and 80% ACN/0.2% TFA, and 2 mg of processed ZIC-HILIC materials was added to the peptide solution and incubated for 2 h at 37°C. Finally, the mixture was transferred to a 200 μl pipette tip packed with a C8 membrane, and washed twice with 80% ACN/0.2% TFA. After enrichment, intact glycopeptides were eluted thrice with 70 μl of 0.1% TFA, and dried using a SpeedVac concentrator. The enriched intact glycopeptides were digested using 1 U PNGase F dissolved in 50 μl of 50 mM NH₄HCO₃ for 2 h at 37°C. The reaction was terminated by adding 0.1% FA. The de-N-glycopeptides and O-glycopeptides were dried using a SpeedVac concentrator for further analysis.
Liquid chromatography-Tandem Mass Spectrometry Analysis

All the samples were analyzed using higher energy collision-induced dissociation (HCD) in mass spectrometry (Orbitrap Fusion Lumos mass spectrometer). In brief, intact O-glycopeptides and de-N-glycopeptides were dissolved in 0.1% FA and separated on a column (ReproSil-Pur C18-AQ, 1.9 μm, 75 μm inner diameter, 20 cm length; Dr Maisch) over a 78 min gradient (buffer A, 0.1% FA in water; buffer B, 0.1% FA in 80% ACN) at a flow rate of 300 nL/min. MS1 was analyzed with a scan range (m/z) of 350–1,550 at an Orbitrap resolution of 120,000. The RF lens, AGC target, maximum injection time, and HCD type were 5.0e4, 80%, 1.0e6, 50 ms, and 15 s, respectively. MS2 was analyzed with an isolation window (m/z) of 2 by an Orbitrap resolution of 15,000. The AGC target, maximum injection time, and HCD type were 5.0e4, 80 ms, and 35%, respectively. For further verification, the same samples were analyzed using electron transfer/higher energy collision-induced dissociation (EThcD) in mass spectrometry (Orbitrap Fusion Lumos mass spectrometer). The MS1 was analyzed with a mass range of 400–1,600 at a resolution of 120,000 at 200 μm/z. The RF Lens was set as 30% and the maximum injection time (MIT) was 100 ms. The MS2 was analyzed in quadrupole mode and the isolation window was 20 μm/z. The EThcD collision energy type was 35%. The MIT was set at 250 ms and cycle time was set at 3 s.

Data analysis

Raw data files were searched against the SARS-CoV-2 S protein sequence using Byonic™ software (version 3.6.0, Protein Metrics, Inc.), with the mass tolerance for precursors and fragment ions set at ±10 and ±20 ppm, respectively. HCD or EThcD was chosen as the fragmentation type. Two missed cleavage sites were subjected to trypsin or trypsin/Glu-C digestion. The fixed modification was carbamidomethyl (C), and the variable modifications included oxidation (M), acetyl (protein N-term), and de-amidation (N). In addition, the six most common O-glycans (HexNAc(1) with mass of 203.079 Da; HexNAc(2) with mass of 406.159 Da; HexNAc(1)Hex(1) with mass of 365.132 Da; HexNAc(2)Hex(1) with mass of 568.212 Da; HexNAc(1)Hex(1)NeuAc(1) with mass of 656.228 Da; and HexNAc(2)Hex(1) with mass of 947.323 Da) were specified as O-glycan modifications for intact O-glycopeptides. We then added the protein database and the decoy database. All other parameters were set to the default values, and protein groups were filtered using a 1% false discovery rate, based on the number of hits obtained for the searches against the databases. Stricter quality control methods for intact O-glycopeptide identification were implemented; they required a score of not less than 300, and at least six amino acids to be identified. Furthermore, all the glycopeptide-spectrum matches (GPSMs) were examined manually by checking the oxonium ions and b/y/c/z ions to ensure the correct identification of the glycopeptides and their glycan compositions, and distinguish the confident O-glycosites and their linked glycans from the uncertain glycosites within a specific glycopeptide. In addition, these O-glycosites had to be identified repeatedly at least twice. Model building based on the Cryo-EM structure (PDB: 6VSB) of the SARS-CoV-2 S protein was performed using PyMOL.

RESULTS AND DISCUSSION

Strategy for Intact O-Glycopeptide Analysis

Our previous study, as well as others, have revealed site-specific N-glycosylation of recombinant S proteins (Watanabe et al., 2020; Zhang et al., 2020). Two or more potential O-glycosites have also been discovered using LC-MS/MS in recent studies (Shajahan et al., 2020b; Bagdonaite et al., 2021). However, comprehensive O-glycosylation analysis of the SARS-CoV-2 S protein has not been performed. In the present study, we aimed to characterize the O-glycosylation landscapes of SARS-CoV-2 recombinant S proteins by analysis of intact O-glycopeptides, including potential O-glycosites and their linked O-glycans.

The strategy for intact O-glycopeptide analysis is shown in Figure 1A. The recombinant SARS-CoV-2 S proteins were digested using trypsin or a mixture of trypsin and Glu-C to cover as many potential O-glycosites as possible. Then, intact glycopeptides were enriched using ZIC-HILIC (Pohlentz et al., 2016), and de-N-glycosylated with PNGase F to avoid interference from non-glycopeptides and N-glycopeptides. Finally, intact O-glycopeptides were analyzed using a high-resolution mass spectrometer, and their mass spectra were characterized using Byonic™ software and validated manually (Zhang et al., 2018). It is worth remarking that O-glycosylation assignment to a specific amino acid by Byonic™ is not always confident when multiple Ser/Thr residues are present within the glycopeptide, especially when using HCD-type MS2 fragmentation. These O-glycosites were classified into the potential sites in this study.

The S protein expressed in insect cells contained 1,209 amino acids (residues 16–1,213), including 94 Thr and 92 Ser residues regarded as potential O-glycosites. The spike protein S1 subunit expressed in human cells contained 681 amino acids (residues 16–685), including 57 Thr and 50 Ser residues as potential O-glycosites (Supplementary Figure S1). Combined digestion strategy can improve glycosite identification and glycoprotein sequence coverage (Chen et al., 2011). To evaluate our method based on MS analysis and data analysis by Byonic in this study, we first analyzed the two previously reported O-glycosites, T323 and S325. The spike protein subunits S1 and S2 expressed on human cells were digested by trypsin and/or chymotrypsin, and analyzed by stepped HCD product triggered CID (HCD-pd-CID) without glycopeptide enrichment and PNGase F digestion (Shajahan et al., 2020b). As shown in Figure 1B, Byonic™ analysis disclosed the presence of the O-glycopeptide 320VQPTESIVR328 with an uncertain O-glycosite at T323 based on the b/y ions with or without glycan retention. It is worth noting that S325 was an alternative glycosite in this peptide because only the “∼y7+203” ion with glycan retention was detected, although Byonic™ tended to assign the T323 according to the y4/y5 ions without the linked glycan. However, the presence of “∼y6∼y7∼y8∼b7” ions without the

Frontiers in Chemistry | www.frontiersin.org 3 September 2021 | Volume 9 | Article 689521

O-glycosylation of SARS-CoV-2 Spike Proteins

Zhang et al.
glycan indicates the fact that the b/y ions produced by HCD tends to lose their glycans (Pap et al., 2018). Without sufficient b/y ions with glycan retention, the confident O-glycosite cannot be determined in the glycopeptide. Shajahan et al. has reported that T323 seems a predominantly occupied site in an O-glycopeptide with the same peptide sequence 320VQPTESIVR328 through HCD fragmentation (Shajahan et al., 2020b), suggesting T323 is a high-probability glycosite. Similarly, the representative HCD-MS/MS spectra in Figure 1C revealed the presence of an uncertain O-glycosite at S325, which also could be T323 due to high frequency loss of entire glycan in HCD fragmentation. Hence, both T323 and S325 are uncertain O-glycosites and could not be confident identified by the HCD-MS/MS spectra in Figures 1B,C. These results indicate that our strategy is feasible for O-glycosylation profiling.

### Comprehensive O-Glycosylation Profiling of Recombinant SARS-CoV-2 S Protein Expressed in Insect Cells

The S protein produced by the baculovirus insect cell expression system contained 186 potential O-glycosites. Using our aforementioned strategy, a total of 255 intact O-glycopeptides composed of 50 peptides backbones and 43 uncertain O-glycosites were discovered by HCD (Supplementary Table S1 and Figure S2). In these glycopeptides, 40 potential O-glycosites, except S477, T572, and T732 were found repeatedly using trypsin alone. Using trypsin combined with Glu-C, three more O-glycosites (S325, T333, and T1066) were missed due to combinational digestion (Figure 2A). Hence, although trypsin digestion can yield good identification results,
trypsin combined with Glu-C digestion should be considered as complementary step because that some suitable glycopeptides can be easily detected by mass spectrometry. Furthermore, we mapped these O-glycosites to the amino sequences, and found that the O-glycosites appeared in several areas, especially in the N- and C-termini of the S protein (Figure 2B). It is notable that the O-glycosites T323, S325, T333, S345, and S477 were located in the receptor-binding domain (RBD). These results indicate that the SARS-CoV-2 S protein is an O-glycoprotein with a large number of O-glycosites. In addition, the number of O-glycosylated Thr residues 25) was higher than that of O-glycosylated Ser residues 18) (Figure 2B). This result is consistent with those of previous studies on O-glycosites (Zhang et al., 2018). Finally, a global O-glycan composition analysis was performed (Figure 2C). Six O-glycan compositions were identified on these sites, including

**FIGURE 2** | Comprehensive O-glycosylation characterization of recombinant SARS-CoV-2 S protein (S1+S2 ECD, His tag) expressed in insect cells. (A) Uncertain O-glycosites identified using trypsin (T) or trypsin/Glu-C (TG) in three replicates. (B) Mapping of identified O-glycosites to amino acid sequences. RBD is highlighted in yellow. (C) O-glycan compositions on each site.
HexNAc(1), HexNAc(2), HexNAc(1)Hex (1), HexNAc(2)Hex (1), HexNAc(1)Hex (1)NeuAc (1), and HexNAc(1)Hex (1)NeuAc (2). Regarding the frequency of these glycans on different glycosites, occupancies with HexNAc(1)Hex (1), HexNAc(1), HexNAc(2)Hex (1), HexNAc(1)Hex (1)NeuAc (1), HexNAc(2), and HexNAc(1)Hex (1)NeuAc (2) compositions were found on 40, 30, 21, 18, 11, and seven glycosites, respectively. Moreover, most glycosites contained at least two types of O-glycans, a majority of which were non-sialylated (Figure 2C). It’s worth noting that NeuAc would rarely found on insect O-glycans, because that insect cells generally lack adequate levels of the glycosyltransferases to synthesize sialylated products, especially if there’s not NeuAc oxonium ion in the MS2 spectrum (Figure 2C). When that happens, the spectrum may be a half-right identification (right peptide with wrong O-glycans) although with high Byonic score. These results indicate the O-glycans appeared on the recombinant SARS-CoV-2 S protein expressed in insect cells.

Comprehensive O-Glycosylation Profiling of Recombinant SARS-CoV-2 S Protein Expressed in Human Cells

The recombinant SARS-CoV-2 S protein S1 subunit produced by the human cell expression system was used for analysis of the
O-glycans, as the O-glycan compositions in insect cells could be different from those in human cells. Using our aforementioned strategy, 407 intact O-glycopeptides composed of 34 peptides backbones and 30 uncertain O-glycosites (20 Thr and 10 Ser residues) were discovered by HCD (Supplementary Table S2 and Figure S3). 24 and 27 uncertain O-glycosites were found repeatedly using trypsin and a mixture of trypsin/Glu-C, respectively. The trypsin combined with Glu-C digestion can increase the number of identified O-glycosites (Figure 3A). The results showed that the two digestion methods were complementary for O-glycosite identification. Furthermore, we mapped these 30 O-glycosites to the amino sequences. We found that the O-glycosites mainly appeared at the S1 subunit and RBD (Figures 3B, C). It is notable that two conserved O-glycosites, T323 and S325, were located in the RBD of the S1 subunit, and may play a critical role in viral binding with hACE2 receptors (Andersen et al., 2020; Hoffmann et al., 2020). A global O-glycan composition analysis of the S1 subunit was performed. O-glycan occupancies with HexNAc(1)Hex (1)NeuAc (2), HexNAc(1)Hex (1)NeuAc (1), HexNAc(2), HexNAc(1), and HexNAc(2)Hex (1) were found on 29, 25, 24, 16, 10, and nine glycosites, respectively. O-glycans on most glycosites were sialylated (Figure 3D). These results indicate the more complex O-glycosylation and the heterogeneity of O-glycan composition.
compositions on the recombinant SARS-CoV-2 S protein expressed in human cells.

**O-Glycosylation Landscapes of S1 Subunits Expressed in Insect and Human Cells**

Based on the above findings, we further compared the O-glycosylation landscapes of the S1 subunits expressed in insect and human cells. 23 uncertain O-glycosites were present in the S1 subunit expressed in insect cells (Figure 4A). In contrast, 30 O-glycosites were present in the S1 subunit expressed in human cells (Figure 4B). In addition, 16 common O-glycosites (T22, T29, S31, T124, T284, T286, S297, T299, T323, S325, T572, T573, S659, S673, T676, and T678) were discovered in the S1 subunits expressed in both insect and human cells, including the two sites, T323 and S325, located in the RBD. Seven and 14 unique O-glycosites were found in the insect and human cell–produced S1 subunits, respectively (Figure 4C). It’s worth noting that HCD can identify the intact O-glycopeptide confidently while usually failing to distinguish a specific O-glycosite from multiple uncertain glycosites within a glycopeptide. More validation experiments were needed to compare the differences between the two S proteins. Furthermore, the number of S1 subunit O-glycosites occupied by each type of O-glycan compositions was very different. Most O-glycosites of the insect cell–produced S1 subunit contained HexNAc(1)Hex (1) and HexNAc(1). On the other hand, most O-glycosites of the human cell–produced S1 subunit contained HexNAc(1)Hex (1)NeuAc (2), HexNAc(1)Hex (1)NeuAc (1), and HexNAc(1)Hex (1) (Figure 4D). These results implied that the O-glycosite and O-glycan compositions varied with the host cell type, which could be taken into account when using the recombinant proteins for vaccine and drug development.
Significant variations in glycosylation analysis. Even so, these methods and data may be useful for the development of vaccines and targeted drugs.

CONCLUSIONS

In this study, we profiled a comprehensive O-glycosylation pattern of SARS-CoV-2 S proteins using optimized experimental procedure and HCD and EThcD mass spectrometry. There are 255 intact O-glycopeptides composed of 50 peptides sequences and 43 uncertain O-glycosites were discovered by HCD in insect cell–expressed S protein, and most of them were non-sialylated. There are three O-glycosites were confidently identified by EThcD. In contrast, in human S protein, 407 intact O-glycopeptides composed of 34 peptides sequences and 30 uncertain O-glycosites were discovered by HCD, 11 O-glycosites were unambiguously assigned by EThcD, and most of them were sialylated. However, the measurement of O-glycosylation occupancy hasn’t been made. Our results revealed that the SARS-CoV-2 S protein was modified by O-glycans, and that the O-glycosite and O-glycan compositions varied with the host cell type.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

AUTHOR CONTRIBUTIONS

HY, MG, JZ, YC, and JC directed and designed research; YZ and WZ wrote the article. HY, MG, JZ, YC, and JC directed and designed research; YZ and WZ wrote the article.

FUNDING

This work was funded by grants from the National Natural Science Foundation of China (31901038), the China Postdoctoral Science Foundation (2020M670063ZX, 2019M653438), Department of Science and Technology of Sichuan Province (2020YFH0029, 2021YJ0479), 1.3.5 Project for Disciplines of Excellence, West China Hospital, Sichuan University (ZYGJ18014), and Chengdu Science and Technology Department Foundation (2020-YF05-00240-SN).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fchem.2021.689521/full#supplementary-material
REFERENCES

Andersen, K. G., Rambaut, A., Lipkin, W. I., Holmes, E. C., and Garry, R. F. (2020). The Proximal Origin of SARS-CoV-2. *Nat. Med.* 26 (4), 450–452. doi:10.1038/s41591-020-02820-9

Bagdonaite, I., Thompson, A. J., Wang, X., Sogard, M., Fougereoux, C., Frank, M., et al. (2021). Site-Specific O-Glycosylation Analysis of SARS-CoV-2 Spike Protein Produced in Insect and Human Cells. *Viruses* 13 (4). doi:10.3390/v13040551

Chen, Y., Cao, J., Yan, G., Lu, H., and Yang, P. (2011). Two-step Protease Digestion and Glycopeptide Capture Approach for Accurate GlycoSite Identification and Glycoprotein Sequence Coverage Improvement. *Talanta* 85 (1), 70–75. doi:10.1016/j.talanta.2011.03.029

Dong, X., Chen, C., Yan, J., Zhang, X., Li, X., and Liang, X. (2021). Comprehensive O-Glycosylation Analysis of the SARS-CoV-2 Spike Protein with Bioinformatic Trp-Arg Materials. *Anal. Chem.* 93 (30), 10444–10452. doi:10.1021/acs.analchem.0c06634

Frese, C. K., Zhou, H., Taus, T., Alletta, A. F. M., Machtler, K., Heck, A. J. R., et al. (2013). Unambiguous Phosphosite Localization Using Electron-Transfer/higher-Energy Collision Dissociation (ET/ECD). *J. Proteome Res.* 12 (3), 1520–1525. doi:10.1021/pr301130k

Grant, O. C., Montgomery, D., Ito, K., and Woods, R. J. (2020). Analysis of the SARS-CoV-2 Spike Protein Glycan Shield Reveals Implications for Immune Recognition. *Sci. Rep.* 10, 1, 14991. doi:10.1038/s41598-020-71748-7

Hoffmann, M., Kleine-Weber, H., Schroeder, S., Kruger, N., Herrler, T., Erichsen, S., et al. (2020). SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* 181 (2), 271–280. e8. doi:10.1016/j.cell.2020.02.052

Hogan, J. M., Pitteri, S. J., Chrisman, P. A., and McLuckey, S. A. (2005). Characterization of Protein Induced Dissociation. Glycopeptide via Electron Transfer Ion/Ion Reactions and Collision-Induced Dissociation. *J. Proteome Res.* 4 (2), 628–632. doi:10.1021/pr049770q

King, S. L., Joshi, H. J., Schjoldager, K. T., Halim, A., Madsen, T. D., Dziegiel, M. H., et al. (2018). Status Report on the High-Throughput Characterization of Complex Glycoprotein Sequence Coverage Improvement. *Talanta* 163, 18–24. doi:10.1016/j.talanta.2017.04.029

Yang, W., Shah, P., Hu, Y., Toghi Eshghi, S., Sun, S., Liu, Y., et al. (2017). Comparison of Enrichment Methods for Intact N- and O-Linked Glycopeptide Libraries. *Anal. Chem.* 89 (6485), 1444–1448. doi:10.1021/acs.analchem.7b03641

Ye, Z., Mao, Y., Clausen, H., and Vakhрушев, S. Y. (2019). Glyco-DIA: A method for Quantitative O-Glycoproteomics with Silico-boosted Glycopeptide Libraries. *Nat. Methods* 16 (9), 902–910. doi:10.1038/s41592-019-0504-x

Zhang, Y., Zhao, W., Mao, Y., Chen, Y., Wang, S., Zhong, Y., et al. (2020). Site-specific N-Glycosylation Characterization of Recombinant SARS-CoV-2 Spike Proteins. *J. Proteomics* 2020, 013276. doi:10.1016/j.jprot.2017.09.014

Zhou, D., Tian, X., Qi, R., Peng, C., and Zhang, W. (2021). Identification of 22 N-Glycosites on Spike Glycoprotein of SARS-CoV-2 and Accessible Surface Glycopeptide Motifs: Implications for Vaccination
and Antibody Therapeutics. *Glycobiology* 31 (1), 69–80. doi:10.1093/glycob/cwaa052

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Zhang, Zhao, Mao, Chen, Zheng, Cao, Zhu, Hu, Gong, Cheng and Yang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.