

# Impact of Expressing Cells on Glycosylation and Glycan of the SARS-CoV-2 Spike Glycoprotein

Yan Wang, Zhen Wu, Wenhua Hu, Piliang Hao, and Shuang Yang\*

Cite This: *ACS Omega* 2021, 6, 15988–15999

Read Online

ACCESS |



Metrics &amp; More

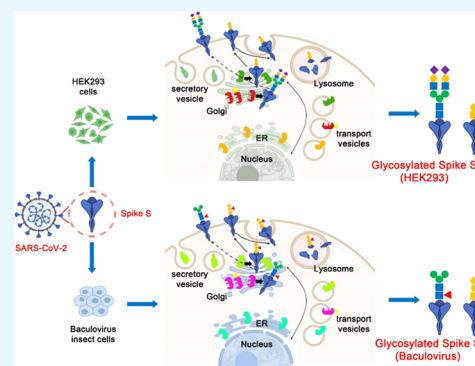


Article Recommendations



Supporting Information

**ABSTRACT:** The spike glycoprotein of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the first point of contact for the virus to recognize and bind to host receptors, is the focus of biomedical research seeking to effectively prevent and treat coronavirus disease (COVID-19). The mass production of spike glycoproteins is usually carried out in different cell systems. Studies have been shown that different expression cell systems alter protein glycosylation of hemagglutinin and neuraminidase in the influenza virus. However, it is not clear whether the cellular system affects the spike protein glycosylation. In this work, we investigated the effect of an expression system on the glycosylation of the spike glycoprotein and its receptor-binding domain. We found that there are significant differences in the glycosylation and glycans attached at each glycosite of the spike glycoprotein obtained from different expression cells. Since glycosylation at the binding site and adjacent amino acids affects the interaction between the spike glycoprotein and the host cell receptor, we recognize that caution should be taken when selecting an expression system to develop inhibitors, antibodies, and vaccines.



## INTRODUCTION

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a strain of novel coronavirus that caused the 2019 pandemic disease (COVID-19). SARS-CoV-2 has close genetic similarity to bat coronavirus. Since its first appearance in Wuhan, China, in December 2019, SARS-CoV-2 has spread globally in a few months.<sup>1</sup> It was confirmed by January 20, 2020, that SARS-CoV-2 can be transmitted from person to person through direct or indirect contact, such as respiratory droplets (coughs or sneezes), airborne, fomite, and urine or feces. As of March 2021, the SARS-CoV-2 virus has caused 2.7 million deaths and 123.2 million cases worldwide.

Similar to the earlier coronavirus strains middle east respiratory syndrome (MERS)-CoV and SARS-CoV that transmits to humans, SARS-CoV-2 consists of four structural proteins, called spike (S), envelope (E), membrane (M), and nucleocapsid (N) proteins. The S, E, and M proteins together form the viral envelope, while the N protein retains the RNA genome.<sup>2,3</sup> The entry of SARS-CoV-2 cells depends on the binding of the viral S protein to the cell receptors and the S protein triggered by the host cell proteases. Studies have shown that the cell entry process engages angiotensin-converting enzyme 2 (ACE2) to bind the S protein and uses transmembrane protease serine 2 (TMPRSS2) to trigger the S protein.<sup>4,5</sup> TMPRSS2 not only cleaves and activates the spike glycoprotein for membrane fusion but also splits ACE2 into two to enhance viral infectivity.<sup>6</sup> Because ACE2 and TMPRSS2 are highly present in the respiratory system, digestive tract, and gastrointestinal tract (such as human

airway epithelium),<sup>7,8</sup> bronchial transient secretory cells,<sup>9</sup> nasal epithelial cells,<sup>10</sup> human ocular surface,<sup>11</sup> and small intestine,<sup>12</sup> various routes of infection may occur when SARS-CoV-2 comes into contact with humans through any of these organs. Furin, another receptor highly expressed in the lungs, binds to the spike and cleaves the furin cleavage site (FCS) of SARS-CoV-2.<sup>13</sup> The presence of ACE2, TMPRSS2, and furin in these cells and tissues may indicate that there are multiple routes of transmission through their respective viral infections.

The spike glycoprotein of coronavirus plays a key role in virus infection, mediates virus entry, and is a primary determinant of cell tropism and pathogenesis.<sup>14</sup> The spike S1 is the first possible point of contact for recognition and binding of host receptors (ACE2, Furin, or GRP78 via CD147<sup>15</sup>), allowing subsequently conformational changes in S2, thereby promoting the fusion between the viral envelope and the host cell membrane. According to reports, the binding affinity of the SARS-CoV-2 S1 receptor-binding domain (RBD) to ACE2 is considerably higher than that of SARS-CoV,<sup>16,17</sup> leading to severe infection and the widespread of the SARS-CoV-2 virus. Although the mortality rate has decreased from ~10% of

Received: April 3, 2021

Accepted: June 2, 2021

Published: June 11, 2021



SARS-CoV to 1–5% of SARS-CoV-2, the number of deaths caused by SARS-CoV-2 is substantially higher than that of SARS-CoV, e.g., over 2.7 million for the former and 812 for the latter globally to date. Therefore, it is important to understand the structure of the spike glycoprotein and the mechanism of infection.

The spike glycoprotein deploys S1 for attachment to the host cell and S2 for fusion. Obviously, the high affinity promotes the attachment of S1 to the host cell and increases the spread of the virus. The detailed structural comparison of S1 between SARS-CoV and SARS-CoV-2 shows that 10 regions in the S1 domain play critical roles in ACE2 binding; mutations in certain amino acid residues in these regions result in low affinity of S1 to ACE2.<sup>4</sup> In contrast, SARS-CoV-2 mutations on some amino acids may help enhance affinity, such as Y442 in SARS-CoV to L457 in SARS-CoV-2, N479 to Q493, Y484 to Q498.<sup>4</sup> Thus, any mutations in amino acid residues or post-translational modification (PTM) of amino acids may affect the attachment of the spike S1 to the host cell receptors. Because the spike is a glycoprotein, its glycosylated variants have a profound effect on the affinity and infectivity of SARS-CoV-2. Recent studies have identified 22 N-glycosites in the protomer of the trimeric spike and have a high-density N-glycan mask on the surface of the viral protein, similar to the S1 subunit of MERS-CoV.<sup>18,19</sup> Several studies have also detected trace levels of O-glycosites at T323 and S325 of the spike glycoprotein<sup>19,20</sup> and T678 near FCS occupied by core-1 and core-2 structures.<sup>21</sup> Recent studies have identified 25 O-glycosites in the S1 of the spike glycoprotein expressed from HEK293 cells, of which 16 O-glycosites are located within the three amino acids from the N-glycosites.<sup>22</sup> These results are consistent with our predictions using ISOGLYP, indicating that S1 RBD is highly O-glycosylated in SARS-CoV-2. On the other hand, as observed in the influenza viruses, when viral glycoproteins are expressed in different cell systems, their glycosylation can change.<sup>23,24</sup> Yet, it has not been studied whether the expression cell system has an impact on the O-glycosylation of SARS-CoV-2 S1 RBD.

In this study, we intend to comprehensively characterize N-linked and O-linked glycosylation of the spike S1 subunit of SARS-CoV and SARS-CoV-2 produced by different expression host cells. We recognize that host expression may alter the glycosylation pattern of spike glycoproteins. HEK293 cells and baculovirus-insect system Hi5 cells are used for virus production and recombinant spike glycoprotein production in our work. The effect of host cell lines on viral protein glycosylation has been reported. The influenza A virus glycoprotein can contain structures of paucimannose (Sf9 cells), core-fucosylated bisected N-GlcNAc (embryonated hen egg), or sialylated biantennary glycans (HEK293).<sup>23</sup> baculovirus-insect cells, already used in influenza and human papillomavirus (HPV), is an ideal baculovirus expression system for the production of recombinant spike glycoproteins and vaccines.<sup>25</sup> baculovirus-insect cells can synthesize glycans with one or two core fucoses. There is a report of glucuronic acid (GlcA) in the cells,<sup>26</sup> even though other insect cells may have GlcA residues.<sup>27</sup> It should be investigated whether baculovirus-insect cells have GlcA and other glycans to analyze the glycosylation of the spike glycoprotein.

To reveal these uncertainties, we compared the S1 subunits of Spike expressed in HEK293 cells and baculovirus-insect Hi5 cells (Table 1). The spike S1 was digested with trypsin, and then glycopeptides were enriched using hydrophilic interaction

**Table 1. Recombinant Spike S1 Expressed in Different Expression Cells<sup>a</sup>**

sample	catalog	description	species	expression host	sequence
BIC1	40150-V08B1	spike S1	SARS-CoV	baculovirus-insect	M7-R667
BIC2	40591-V08B1	spike S1	SARS-CoV-2	baculovirus-insect	V16-R685
HEK2	40591-V08H	spike S1	SARS-CoV-2	HEK293	V16-R685

<sup>a</sup>Samples were purchased from Sino Biological.

liquid chromatography (HILIC). The enriched glycopeptides were analyzed by liquid chromatography–mass spectrometry (LC-MS/MS) using electron-transfer/higher-energy collision dissociation (ETHcD) fragments. In another experiment, N-glycans and O-glycans were released from spike S1 and evaluated using a Bruker Autoflex Matrix-Assisted Laser Desorption/Ionization (MALDI)-MS.

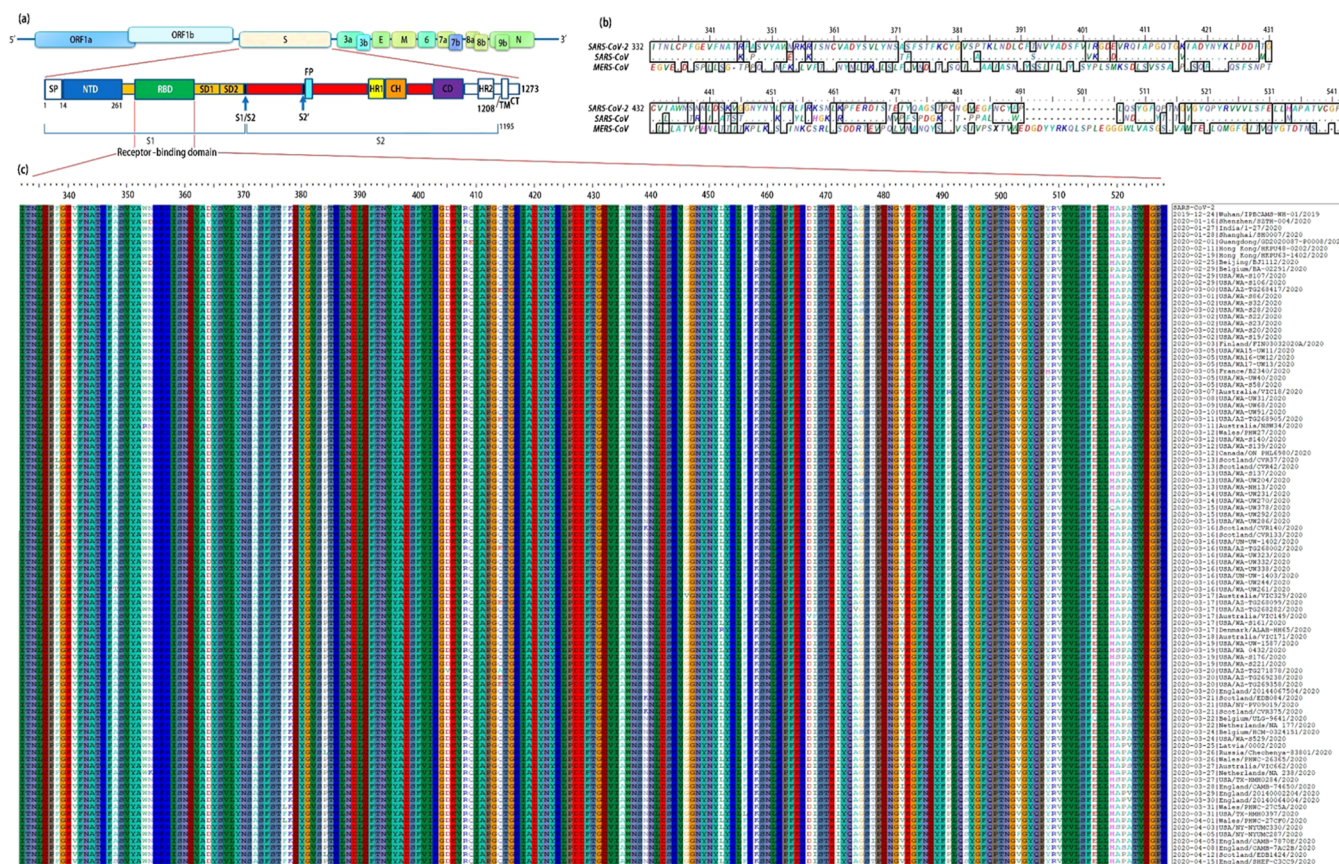
## RESULTS AND DISCUSSION

**Most Diverse Mutations of Amino Acids Occurred in the S1 Domain of the Spike Glycoprotein.** The global initiative on sharing all influenza data (GISAID) has updated the SARS-CoV-2 genome and the spike glycoprotein sequence based on data submitted by laboratories and research institutes around the world. As of February 2021, we have downloaded more than 200 000 protein sequences of the spike glycoproteins. After removing redundant and incomplete sequences, we found that there are 98 unique spike glycoproteins, most of which have mutations in the receptor-binding domain (RBD) of spike S1 (Figure 1). The sequences are arranged according to their submission date (strain list is given in Table S1). Figure 1a illustrates the schematic structure of SARS-CoV-2 and its spike glycoprotein, and Figure 1b compares the alignment of SARS-CoV-2, SARS-CoV, and MERS-CoV. Genetic analysis showed 79% similarity between SARS-CoV and SARS-CoV-2, and the amino acid sequence identity was 76.47%;<sup>28</sup> the sequence alignment between MERS-CoV and SARS-CoV showed significant differences.<sup>29</sup> There were 51 amino acid changes between SARS-CoV and SARS-CoV-2, or 25.8% variation. Importantly, the variation falls in several sites that are critical for binding affinity to the host cell receptors.<sup>4</sup> From 12/2019 to 05/2020, amino acid mutations were observed at 19.3% positions within the RBD domain (Figure 1c). This result indicates that the diversity of SARS-CoV-2 is caused by its frequent mutation on the spike RBD. Thus, it is essential to clarify the spike RBD domain variation to provide necessary information for the development of inhibitors, antibodies, and vaccines.

**N-Glycosylation of SARS-CoV-2 Regulated by the Expression System.** The purified recombinant S1 proteins expressed in HEK293 cells (HEK2) and baculovirus-insect cells (BIC2 and BIC1) (Table 1) were purchased from Sino Biological. The analysis of each sample was performed in triplicate. Each N-glycosite was plotted using relative abundance related to all N-glycosites.

The expression system impacts N-glycosylation and the types of N-glycans at each site. As shown in Figure 2a,b, the spike S1 expressed in HEK293 cells has 12 N-glycosites. When expressed in baculovirus-insect cells, it will carry an additional N-glycosite N603. N-glycans show distinct patterns between the proteins expressed by HEK2 and BIC2. For example, N17





**Figure 1.** Amino acid mutation predominantly occurred on the receptor-binding domain of the SARS-CoV-2 spike glycoprotein from 12/2019 to 05/2020. (a) Domains of SARS-CoV-2 virion include ORF1a&b, spike (S), 3a, 3b, envelope (E), membrane (M), 6, 7a, 7b, 8a, 8b, 9b, and nucleocapsid (N). The spike S1 domain consists of an N-terminal domain (NTD), a receptor-binding domain (RBD), a subdomain 1 (SD1), and a subdomain 2 (SD2); the other domains are S2, heptad repeat 1 (HR1), central helix (CH), connector domain (CD), HR2, transmembrane (TM), and cytoplasmic tail (CT). S1/S2 is the protease cleavage site, FP is the fusion peptide, and S2' is the protease cleavage site. (b) Spike RBD sequence alignment between SARS-CoV-2, SARS-CoV, and MERS-CoV. (c) Alignment on RBD of SARS-CoV-2 strains from 12/2019 to 05/2020. The 98 complete and unique sequences are listed, most of which are conserved. The amino acid mutations are highlighted with white bars, while few mutations are observed in other domains of spike glycoproteins.

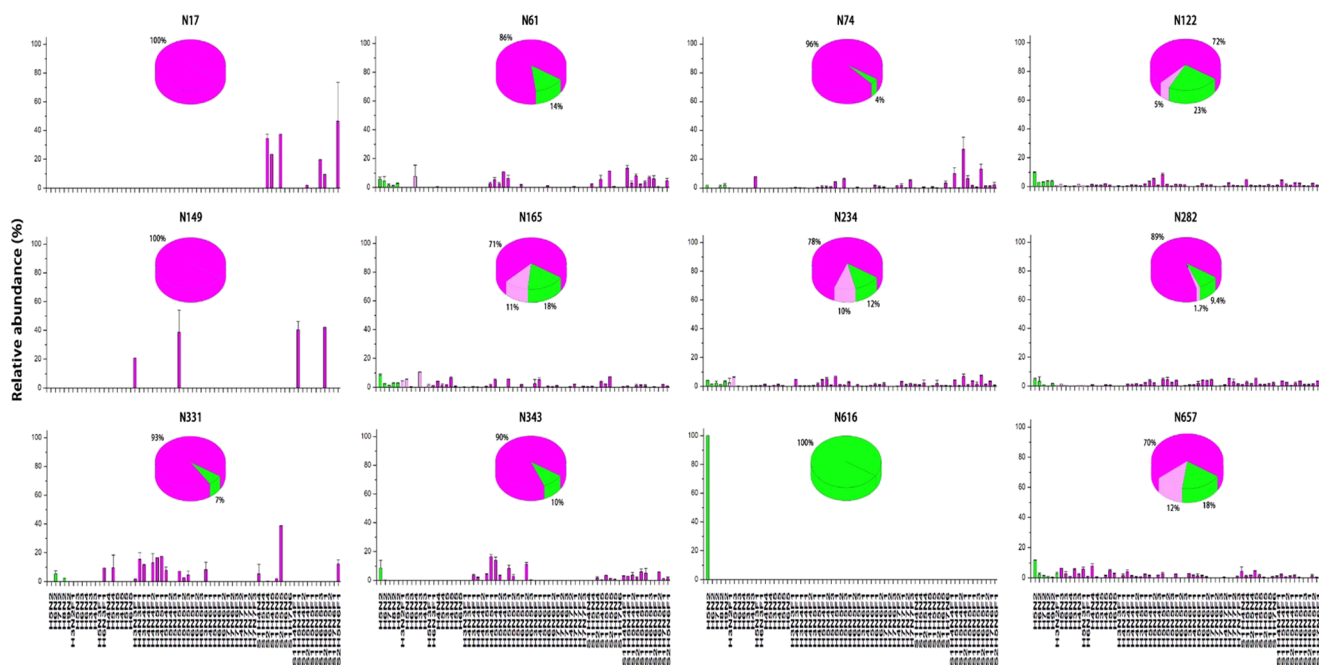
only exhibits complex N-glycans in HEK2, and N17 in BIC2 predominantly contains complex N-glycans with 4% high mannose. A similar observation was also found in N149 of HEK2. On the other hand, HEK2 N616 only has the Man5 (Man5GlcNAc2: Man = Mannose, GlcNAc = N-acetylglucosamine), while N616 from BIC2 mainly contains complex N-glycans, a small amount of hybrid and high-mannose N-glycans. These results indicate that the N616 site from HEK2 cells can be accessed by  $\alpha$ 1,2-mannosidases, but not as much as GlcNAcT-I.<sup>18</sup> Other sites containing complex and high-mannose N-glycans, such as N61, N74, N331, and N343 in HEK2, or N74, N234, N282, and N331 in BIC2, are good substrates for GlcNAcT-I when forming complex N-glycans. N122, N165, N234, N282, and N657 in HEK2 show hybrid N-glycans; N-61, N122, N149, N165, N603, N616, and N657 in BIC2 also have hybrid N-glycans, indicating that the N-glycan process of SARS-CoV-2 depends on the expression system. Moreover, the sialylation distribution of N-glycans is strikingly different between HEK2 and BIC2. Except for N616, all other N-glycosites in HEK2 contain large amounts of sialylated N-glycans. Further linkage analysis by matrix-assisted laser desorption ionization-MS (MALDI-MS) showed that these sialic acids have  $\alpha$ 2,3 or  $\alpha$ 2,6 linkage (Figure 2d and Table S2a),<sup>30</sup> suggesting that these peptide substrates may be

processed by sialyltransferases (e.g., ST3Gal4 or ST6Gal1). Our results are consistent with previous studies on the SARS-CoV-2 spike proteins recombinantly expressed on the HEK293 supernatant,<sup>20,31</sup> except for the identification of N-glycans in N17 and N603 in our study, even though the number of N-glycans observed in these N-glycosites is limited.

SARS-CoV expressed in baculovirus-insect (BIC1) has 14 N-glycosites, and SARS-CoV-2 expressed in baculovirus-insect (BIC2) has 13 N-glycosites. BIC1 and BIC2 produce high-mannose and complex N-glycans, and all of these N-glycans contain fucosylated complex types such as Man3GlcNAc2-Fuc1, or known as paucimannose specific to the insect. These results demonstrate the synthesis of core fucose in the presence of  $\alpha$ 1,3-fucosyltransferase in the baculovirus-insect cells.<sup>32</sup> Conversely, almost no sialylated N-glycans were identified in the BIC1 or BIC2, although treatment of baculovirus-insect with a  $\beta$ -N-acetylglucosaminidase inhibitor may produce terminally sialylated N-glycans.<sup>33</sup> N-glycosylation primarily glycosylated by high-mannoses is located at N61, N122, and N234 in BIC2 and N65, N227, and N318 in BIC1. The subtle difference in N-glycosylation may be attributed to the change in the amino acid sequence between BIC2 and BIC1. Generally, the N-glycan profile is highly conserved between BIC2 and BIC1 (Figure 2d).

(a) HEK293

High-mannose Hybrid Complex (fucose, sialic acids etc.)



(b) Baculovirus-insect (BIC2)

High-mannose Hybrid Complex (fucose, sialic acids etc.)

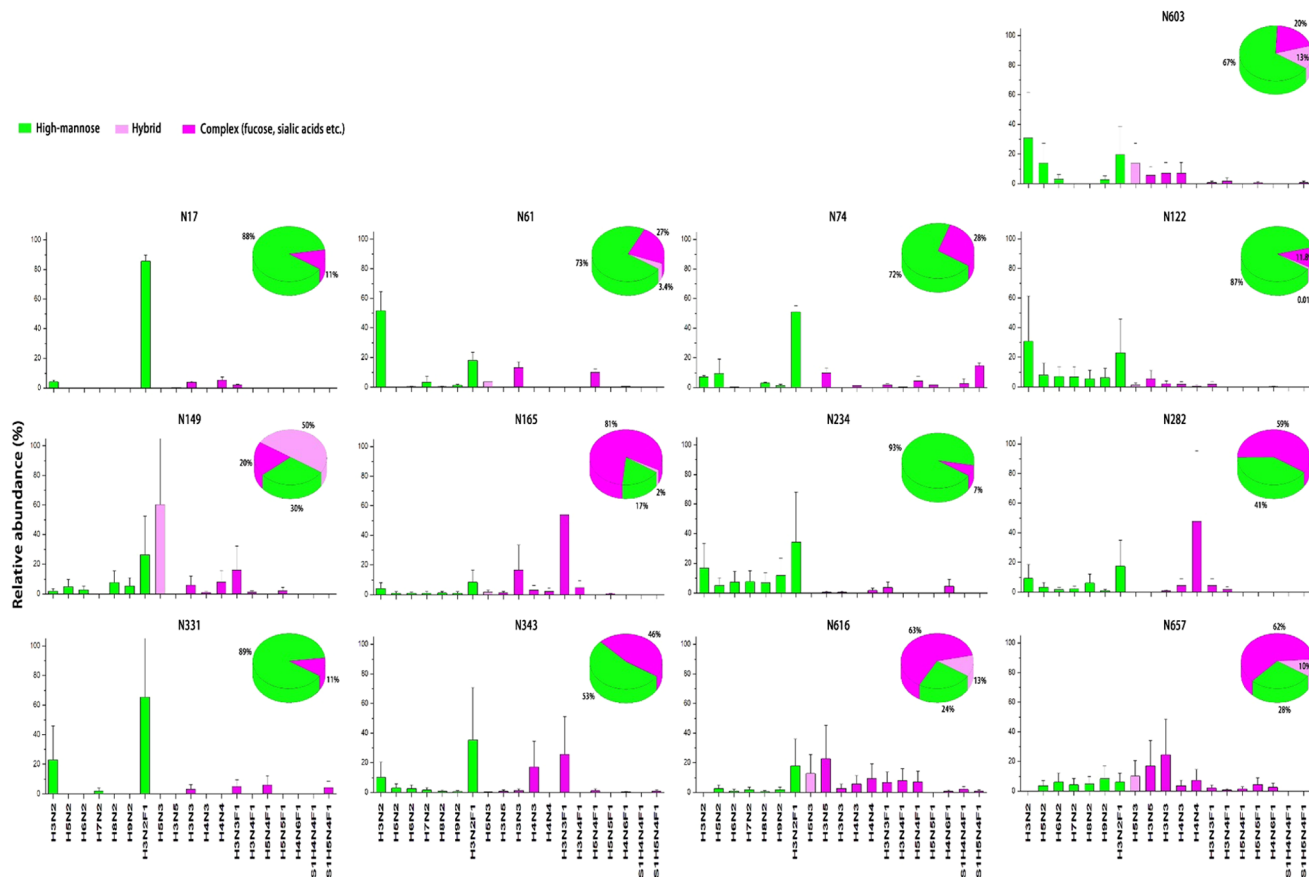
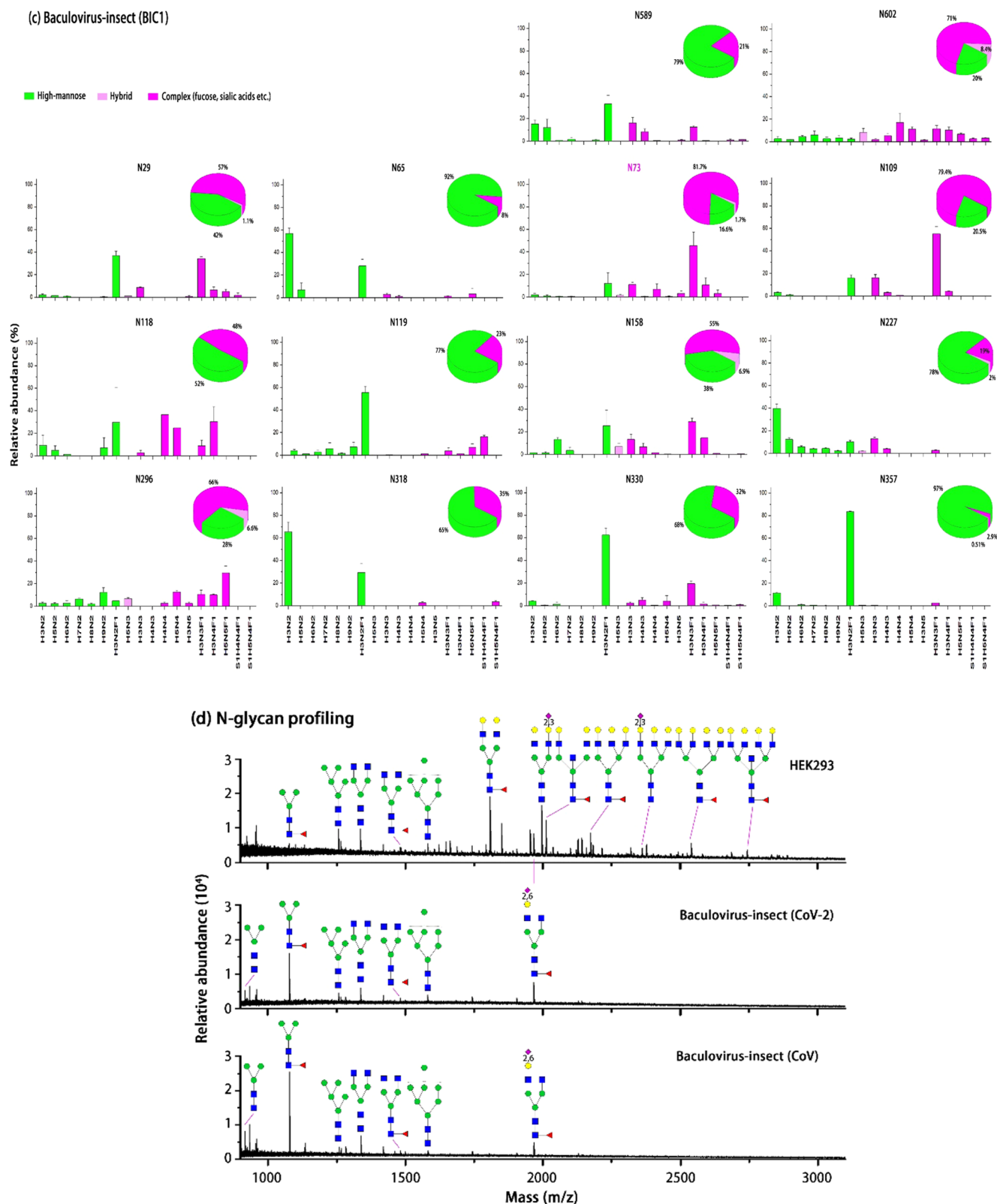


Figure 2. continued

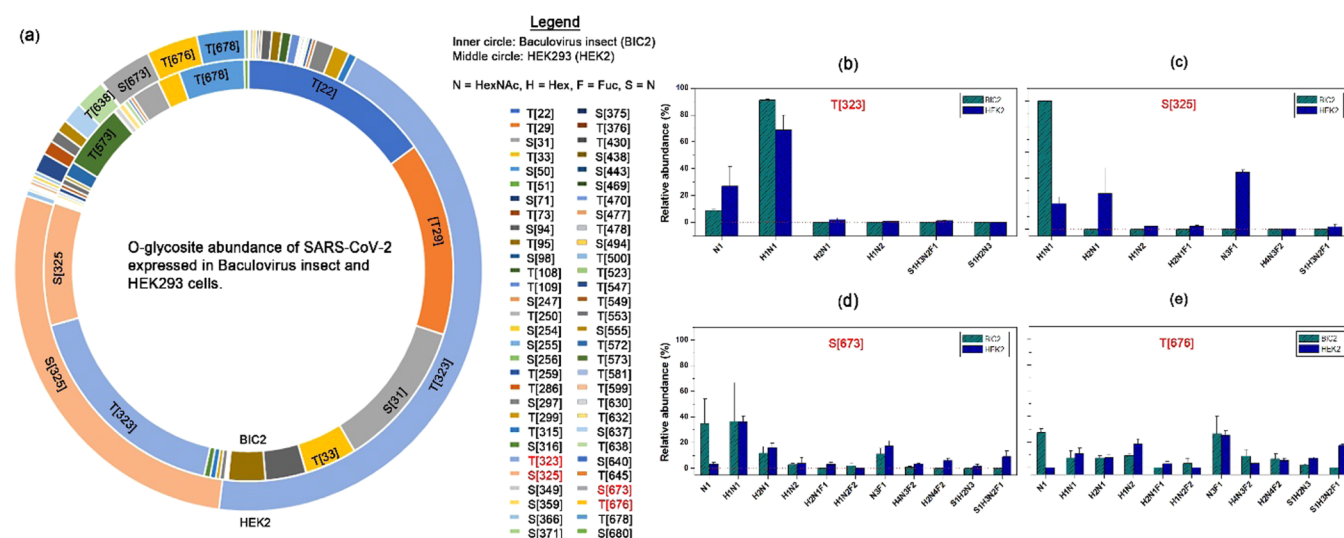




**Figure 2.** Site-specific characterization of N-glycosylation of the S1 domain of SARS-CoV and SARS-CoV-2 spike glycoprotein. (a) SARS-CoV-2 virus expressed in HEK293 cells. Twelve N-glycosites in S1 were identified by LC-MS/MS. N-glycans are divided into high-mannose (green), hybrid (light purple), and complex (purple). N-glycosites, N17 and N149, are attached by complex N-glycans, N616 only has high-mannose (Man5), and other sites are predominantly complex types. Among these sites, N165, N234, and N657 have more than 10% hybrid N-glycans. (b) SARS-CoV-2 virus expressed in baculovirus-insect. In addition to 12 N-glycosites similar to HEK293 cells, another N-glycosite N603 was detected. High-mannose and complex N-glycans are present in all N-glycosites, while hybrid N-glycans are present in N61, N122, N149, N165, N343, N616, N657, and N603. (c) SARS-CoV virus expressed in baculovirus-insect cells. There are 14 N-glycosites in SARS-CoV. High-mannoses are predominantly present in N65, N227, and N318. Complex N-glycans are highly abundant in N29, N73, N109, N118, N119, N158, N296, N330,

Figure 2. continued

N357, N589, and N602. (d) MALDI-MS profiling of N-glycans released from SARS-CoV and SARS-CoV-2. Spike S1 was immobilized on AminoLink plus resins and derivatized by ethyl esterification/ethylenediamine amidation. The most abundant N-glycans are represented, and complete N-glycans for HEK293, CoV-2, and CoV are listed in Table S2. Data are given as mean  $\pm$  standard deviation.



**Figure 3.** Differential O-glycosylation in spike S1 expressed in baculovirus-insect and HEK293. (a) Relative abundance of O-glycosites identified in the Spike S1 domain. The most abundance O-glycosites are labeled in the ring, and the complete list of all O-glycosites are described in the legend. (b) Most abundant O-glycosite, T323, is present in both BIC2 and HEK2. This O-glycosite consists of GalNAc (N1) and GalGalNAc (H1N1). (c) S325 in BIC2 is mainly H1N1, while S325 in HEK2 is more diverse. O-glycosites (d) S673 and (e) T676 reveal more diverse O-glycans in HEK2, including several sialylated species.

**Differential Pattern of O-Glycosites of SARS-CoV-2 in Host Cells.** Table S3 shows the potential O-glycosylation on the SARS-CoV-2 spike glycoprotein predicted by ISO-GlyP.<sup>34,35</sup> In this study, T or S sites marked as “high” were reported in the literature and detected in our work, and our method also detected other O-glycosites marked as “medium” (Table S3). It is worth noting that the detected O-glycosites are mainly in the peptide substrate cluster, e.g., T22, T29, S31, and T33 are in the peptide cluster of T[22]QLPPAYT[29]-NS[31]FT[33]R. This is consistent with the finding that an amino acid substrate containing P (proline) is beneficial to GalNAcTs (UDP-GalNAc:polypeptide *N*-acetylgalactosaminyltransferases (E.C. E.C. 2.4.1.41)) accessible to the T or S residues.<sup>36,37</sup> The charge state surrounding T or S may be a factor because the polarity of the GalNAcTs lectin domain affects glycosylation.<sup>38</sup>

According to N-glycopeptide analysis, spike S1 also showed different O-glycosylation expressed in HEK293 and baculovirus-insect cells. In the HEK293 cells, T323 and T325 are O-glycosylated by GalNAc and GalGalNAc mucin-type O-glycans. S637, T676, and T638 are more abundant than BIC2. In BIC2, T22, T29, S31, T33, S94, T95, T323, and T325 are the most abundant O-glycosites; T572 and T573 are only present in BIC2. These results may imply that the types of GalNAcTs are different in HEK2 and BIC2 because the glycopeptide substrate preferences of GalNAcTs may cause distinct O-glycosylation.<sup>39</sup> It is expected that for the same peptide substrates, such as S323, S325, T676, and T678, there will be some O-glycosites with similar glycosylation. A comparison of the site-specific O-glycan profiles on these O-glycosites is given in Figure 3b–e. We noticed that T323 has O-glycans similar to GalNAc (N1) and GalGalNAc (H1N1).

The other three main O-glycosites have divergent O-glycans. For example, BIC2 has H1N1 at S325, N1, H1N1, H2N1, and H3F1 (F = Fucose) at S673, H1, H1N1, H2N1, H1N2, H1N2F2, H3F1, H4N3F2, and H2N4F2 at T676; HEK2 has H1N1, H2N1, and N3F1 at S325, H1N1, H2N1, H3F1, H2N4F2, and S1H3N2F1 (S = NeuAc) at S673, H1N1, H2N1, H1N2, N3F1, H2N4F2, S1N2N3, and S1H3N2F1 at T676. This demonstrates the combination of the availability of branched glycoenzyme and the preference for GalNAcTs on peptide substrates.<sup>40</sup>

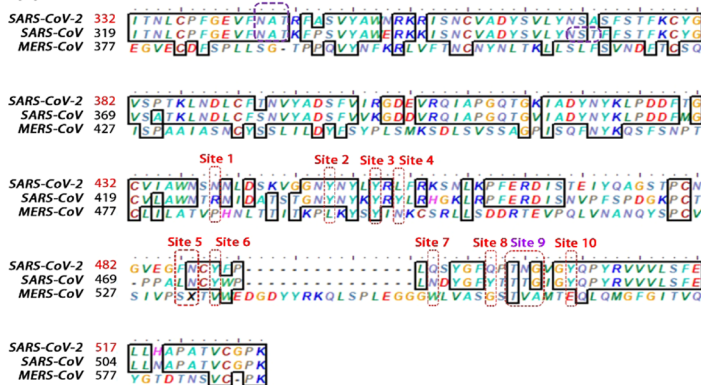
**O-Glycosylation in the BIC2 RBD Domain.** The location of the O-glycosites is different between HEK293 and baculovirus-insect (Figure 3a). We paid special attention to the RBD domain of SARS-CoV-2 expressed in two cell lines. RBD has 197 amino acids starting from I332 to K528. When SARS-CoV-2 was expressed in HEK293 cells, nine O-glycosites were found including S366, T371, T430, S438, S443, S477, T478, S494, and T500 (Figure 4). GalNAc, GalGalNAc, and GalGalNAc2 are the main O-glycans in all O-glycosites, and the abundance of S494 and T500 is high (the area inside the circle in HEK293 represents the relative abundance). These two O-glycosites and T438 are the key positions that may affect the binding affinity of RBD to the ACE2 receptor.<sup>4</sup>

Compared with HEK293 cells, the O-glycosylation of SARS-CoV-2 S1 RBD domain expressed in Baculovirus-insect cells is more diverse and complex. Besides the nine O-glycosites identified in HEK2, six additional O-glycosites were found in BIC2, revealing that the density of O-glycosites is higher in BIC2. Additionally, no sialylated O-glycans were found in BIC2, while HEK2 showed sialic acid at S371, T430, S438, T478, S494, and T500. This is consistent with previous reports that insect cells lack sialyltransferases, rarely produce sialylated

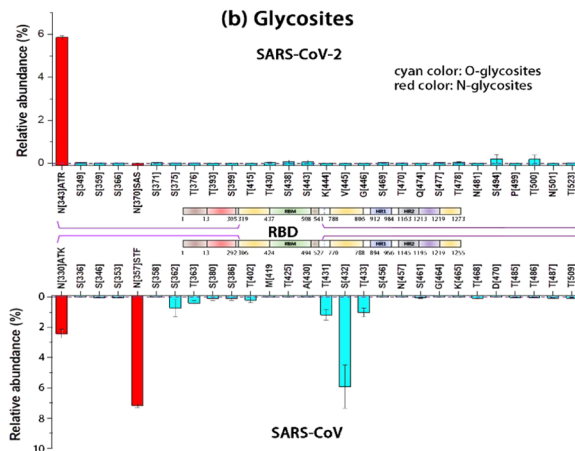




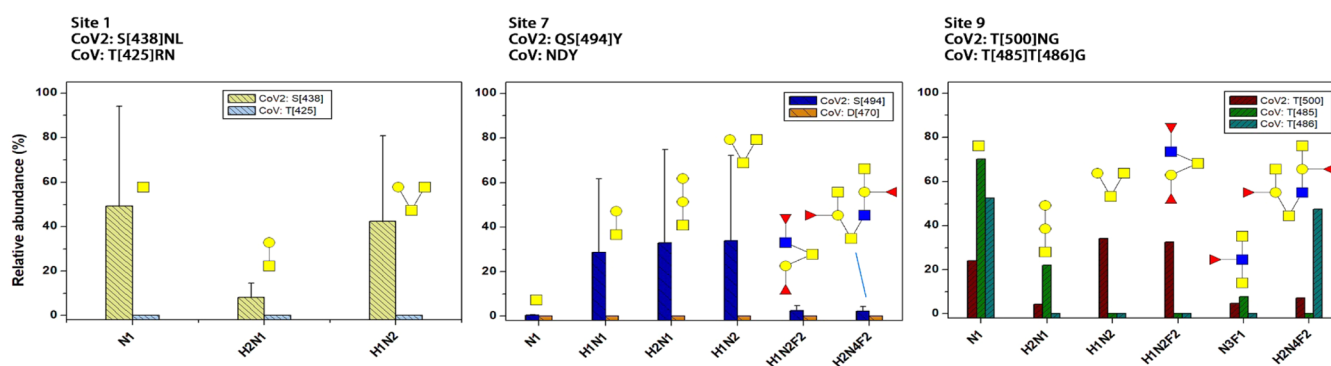
## (a) CoV2 vs CoV and MERS



## (b) Glycosites



## (c) Binding site O-glycosylation



**Figure 5.** O-glycosites in or nearby key ACE2–RBD binding sites. (a) Ten binding sites that are crucial in the ACE2–RBM interaction. These sites are aligned for SARS-CoV-2, SARS-CoV, and MERS. (b) N- and O-glycosites in the RBD domain of the SARS-CoV-2 and SARS-CoV. The red bar is the relative abundance of N-glycosites, and the cyan bar is that of O-glycosites. Each amino acid is aligned based on the sequence described in (a). (c) Three ACE2–RBM binding sites (1, 7, and 9) overlapping with O-glycosites. SARS-CoV-2 has S438, S494, and T500; SARS-CoV has T485 and T486. The RBM, receptor-binding motif, starts from S438 to Q506.

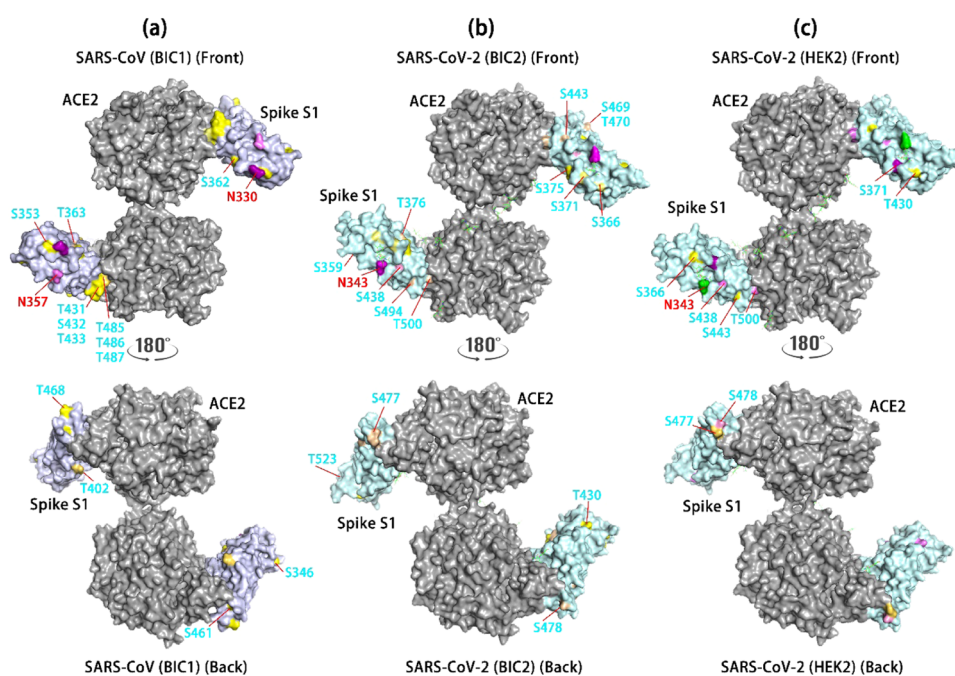
fusion. The affinity between S1 and host cell receptors plays a critical role in viral infection and transmission. The receptor-binding domain of spike S1 has a specific receptor-binding motif (RBM), which may directly interact with the receptor through hydrogen bonds and salt bridges.<sup>45</sup> From S438 to Q506, the RBM domain has 10 sites that directly interact with the ACE2 receptor. The binding kinetics between RBM and ACE2 receptor may be affected by glycosylation on these two proteins,<sup>50</sup> which has been similarly manifested by influenza A virus hemagglutinin<sup>51</sup> and HIV-1 whose encapsulated glycan moieties determine viral propagation.<sup>52</sup> The glycosylation of spike S depends on the host cell line, which can express varying glycoenzymes and transporters, resulting in specificity and heterogeneity.<sup>53</sup> Differential glycosylation not only impacts the infectivity of the virus but also changes the clinical effectiveness of therapeutic products. Thus, we intend to explore how the expression system regulates the glycosylation of spike S1 RBM and secondary structure and compare the glycosylation distribution between SARS-CoV and SARS-CoV-2.

HEK293 and baculovirus-insect cell expression system is used for non-mRNA COVID-19 vaccine development.<sup>54,55</sup> Our results show that the expression cell determines glycosylation of the spike S1 and the type of attached glycans. SARS-CoV-2 derived from baculovirus-insect cells contains high-mannose and fucosylated complex N-glycans and fucosylated mucin-type O-glycans. SARS-CoV-2 in HEK293

cells constructs hybrid and sialylated complex N-glycans and sialylated O-glycans. MALDI-MS analysis found that SARS-CoV-2 in HEK293 contains  $\alpha$ 2,3- and  $\alpha$ 2,6-linked sialic acids. These observations are consistent with the glycan biosynthesis of the expression system. The known glycan biosynthetic pathways of insects can form Man3GlcNAc2Fuc through GlcNAcMan5GlcNAc2 with  $\alpha$ -mannosidase II, core  $\alpha$ 1,3-fucosyltransferase, and N-Acetylglucosaminidase. Complex glycans are further extended by additional glycoenzymes.<sup>56</sup> HEK293 follows the general mammalian glycosylation pathways, forming biantennary, triantennary, or tetraantennary complex glycans in the presence of sialic acid or fucose residues.<sup>57</sup> As expected, we found that SARS-CoV-2 expressed by HEK293 has bisected, fucosylated, and sialylated N-glycans and fucosylated/sialylated O-glycans. When using the same expression host cell, similar glycosylation was still detected in BIC2 and BIC1 despite the different strains.

Glycosite mapping of spike S1 suggests the potential influence of host cells on the binding affinity to the ACE2 receptor. Eight O-glycosites in the RBM domain were identified in the baculovirus-insect and six O-glycosites in HEK293. The difference in glycosylation and the three-dimensional (3D) conformation of spike S1 can improve the interaction with the ACE2 receptor. It is very important to systematically study glycosylation, since the RBM (especially RBM) in the SARS-CoV-2 spike glycoprotein may be the





**Figure 6.** Mapping glycosites of the Spike S1 RBD domain and its human receptor ACE2. N-glycosites are labeled in red and O-glycosites in cyan. The site mapping color represents different types of glycans: yellow = Gal (H1), GalGalNAc (H1N1) without or with minimal fucosylation or sialylation; light yellow = H1 or H1N1 with fucosylation or sialylation; pink = fucosylation and/or sialylation; green = high-mannose; purple = sialylated complex N-glycans; light purple = other types of complex N-glycans. SARS-CoV is based on 3D0H<sup>47</sup> and SARS-CoV-2 on 6WV1.<sup>46</sup> (a) SARS-CoV Spike S1 RBD domain glycosites include T485, T486, and T487 near or within the binding sites between ACE2 and RBD. These sites are H1 and H1N1. The front and back sides of the S1 are illustrated for glycosites. (b) Glycosites on the SARS-CoV-2 spike S1 RBD domain expressed in baculovirus-insect cells. (c) Glycosites on the SARS-CoV-2 spike S1 RBD domain expressed in HEK293 cells.

target for the development of virus attachment inhibitors, neutralizing antibodies, and vaccines.<sup>58</sup> Given that SARS-CoV-2 can be infected and transmitted through many media (lungs, oral, eyes, intestine, etc.), consideration should be given to selecting suitable host cell lines for diagnostic applications and the development of inhibitors, antibodies, or vaccines.

## METHODS

### Sample Preparation of SARS-CoV and SARS-CoV-2.

Recombinant spike S1 was purchased from Sino Biological (HEK2, BIC2, and BIC1) (Table 1). The amino acid sequences of HEK2 and BIC2 were from V16 to K685 and that of BIC1 from M1 to R667. The sample preparation followed the procedure described in Figure S1. Each sample was performed in technical triplicate. First, 40  $\mu$ g of the protein was denatured in high-performance liquid chromatography (HPLC) water at 90 °C/10 min, and half of which was reduced in 12 mM tris(2-carboxyethyl)phosphine hydrochloride (TCEP)/37 °C/1 h and alkylated in 16 mM iodoacetamide (IAA)/room temperature/1 h. The sample was then digested with trypsin (1:25) (Promega, Madison, WI) at 37 °C/overnight. The digest solution was acidified with 30  $\mu$ L of 100% trifluoroacetic acid (TFA) prior to solid-phase extraction (SPE) cartridge C18 cleanup (Waters, Milford, MA). An in-house packed Amide-80 (Tosoh Bioscience LLC, King of Prussia, PA) HILIC SPE column was used to further enrich glycopeptides.<sup>59</sup> The glycopeptides and flow-through peptides after HILIC were analyzed using LC-MS/MS.

The remaining 20  $\mu$ g of the protein after denaturation was conjugated with an Aminolink plus coupling resin (Thermo Fisher Scientific, Waltham, MA) for glycan analysis. The solid-phase method is called glycoprotein immobilization for glycan

extraction (GIG),<sup>60</sup> in which  $\alpha$ 2,6-linked sialic acid underwent an ethyl esterification reaction (0.5 M *N*-(3-dimethylamino-propyl)-*N'*-ethylcarbodiimide hydrochloride (EDC-HCl) and 0.5 M 1-hydroxybenzotriazole hydrate (HBT), 250  $\mu$ L) and  $\alpha$ 2,3-linked sialic acid through a carbodiimide coupling (1 M *p*-toluidine in the presence of EDC (pH 4–6)).<sup>59</sup> First, we used 1  $\mu$ L of PNGase F (New England BioLabs, Ipswich, MA) in 25 mM ammonium bicarbonate to release N-glycans; the remaining sample on the Aminolink resin was further processed to release O-glycans through  $\beta$ -elimination (0.1 M NaOH) and permethylation. The permethylated O-glycans were purified using a C18 SPE cartridge and eluted with 300  $\mu$ L of 60% ACN in 0.1% TFA. Glycans were analyzed by MALDI-time-of-flight/TOF-MS (MALDI-TOF/TOF-MS) (Bruker Autoflex).

**MALDI-TOF/TOF-MS Identification of Glycans.** The eluted glycans in 60% ACN (0.1% TFA) were spotted onto a  $\mu$ Focus MALDI plate (384 circles; Hudson Surface Technology, West New York, NJ), together with 1  $\mu$ L of 10 mg/mL dihydroxybenzoic acid (DHB) matrix in the presence of 2% *N,N*-dimethylaniline (DMA) (50% ACN in 0.1 mM NaCl). The plate was dried on the top of a 50–60 °C hot plate. Each MALDI-MS test was performed in triplicate for 8000 shots. The mass (*m/z*) was searched against the glycan database in GlycoWorkBench.<sup>61</sup> For N-glycans, the mass range was set between 900 and 6000 Da, while it was set between 300 and 3000 Da for O-glycans.

**LC-MS/MS Analysis of Glycopeptides.** The samples were analyzed using a Dionex U3000 nanoHPLC system connected to a Thermo Orbitrap Fusion Lumos mass spectrometer (Thermo Fisher Scientific). Glycopeptides (1  $\mu$ g) were injected and desalted with an Acclaim PepMap C18

Nano trap column (3  $\mu\text{m}$ , 100  $\text{\AA}$ , 75  $\mu\text{m} \times 2 \text{ cm}$ ) at 5  $\mu\text{L}/\text{min}$  with 100% solvent A (0.1% formic acid in HPLC water) for 5 min. Then, glycopeptides were separated by an Acclaim PepMap 100 nano column (3  $\mu\text{m}$ , 100  $\text{\AA}$ , 75  $\mu\text{m} \times 250 \text{ mm}$ ) using a linear gradient of 2.5–37.5% solvent B (80% ACN, 0.1% formic acid) over 85 min, with a wash at 90% B for 5 min. The column was equilibrated at 2.5% B for 10 min before the next injection. Data-dependent analysis (DDA) was carried out with a duty cycle of 2 s. Precursor masses were detected in the orbitrap at a resolution ( $R$ ) of 120 000 (at  $m/z$  200) with internal calibration (Easy IC). Stepped HCD spectra (HCD energy at 15, 25, and 35%) were acquired for precursors with charges between 2 and 8 and intensities over  $5.0 \times 10^4$  at  $R = 30\,000$ . Dynamic exclusion was set at 20 s. When at least one glycan oxonium fragment ion ( $m/z$  138.0545, 204.0867, 366.1396 Da) was observed within the top 20 most abundant fragments and within 15 ppm mass accuracy, an EThcD spectrum was acquired in the orbitrap at  $R = 30\,000$ . The electron-transfer dissociation (ETD) reagent target was  $2.0 \times 10^5$ , with supplemental collision energy at 15%. The ETD reaction time was dependent on the precursor charge state: 125 ms (ETD reaction time) for charge 2, 100 ms for 3, 75 ms for 4, and 50 ms for  $\geq 5$ .

**Data Analysis.** Through precursor and MS/MS fragmentation matching, the glycan composition analysis was performed in GlycoWorkBench, which uses glycan databases from a consortium for functional glycomics (CFG), Carbbank, GlycomeDB, and Glycosciences. The derivatization of the sialic acid linkages added a mass tag to its residues, namely, 28.031301 on  $\alpha 2,6$ -link or 42.058183 on  $\alpha 2,3$ -link. The identified N-glycans and O-glycans were used as the glycan database for glycopeptide analysis (Tables S2 and S4).

MS/MS spectra were searched using Byos (Protein Metrics, San Carlos, CA) against a spiked protein database compiled in-house. The identified glycans in the MALDI-TOF were used as the glycan database. Search parameters include precursor mass tolerance (15 ppm), HCD fragment mass tolerance (20 ppm), EThcD fragment mass tolerance (20 ppm), missed cleavage (3), oxidation (+15.994915, variable), carbamidomethyl (+57.021464, fixed), common modification ( $\leq 2$ ), rare modification (1), maximum precursor mass (30 000), protein FDR (2%), and missed cleavage (3). The identified glycopeptides were manually verified according to oxonium ions, pep-HexNAc, and  $y$  and  $b$  ions with fragments surrounding an O-glycosite. An example of glycopeptide tandem MS is shown in Figure S2. For a peptide that has multiple glycosites, such as N-glycosite and T/S O-glycosites, we use a fragment ion calculator (<http://db.systemsbio.net:8080/peptomicsToolkit/FragIonServlet.html>) to check the fragmentation mass of glycopeptides.

The quantification of glycopeptides was performed as follows. After searching the LC-MS/MS spectra against Byonic, Byologic further analyzed the Byonic output files. The total area under the curve (AUC) of each glycopeptide was extracted from LC-MS/MS by Byologics. The AUC of the same glycopeptide was summed up, and the relative abundance was estimated by dividing the AUC (single glycopeptide) by the total AUC (all glycopeptides). To quantify glycans on each glycosite, we used the AUC of each glycoform divided by the total AUC of all glycoforms.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acsomega.1c01785>.

List of 98 unique sequences of the SARS-CoV-2 spike receptor-binding domain (Table S1); quantitative profiling of N-glycans from SARS-CoV and SARS-CoV-2 expressed in HEK293 cells and baculovirus-insect cells (Table S2); O-glycosite prediction based on the amino acid sequence of SARS-CoV-2 using ISOGLYP (Table S3); O-glycans for Byonic O-glycopeptide search, O-glycans were prepared by the solid-phase glycan extraction, including  $\beta$ -elimination and permethylation (Table S4); schematic workflow of spike glycoprotein analysis (Figure S1); typical HCD fragmentation of the O-glycopeptide precursor (Figure S2); and MALDI-MS of O-glycans of SARS-CoV-2 in BIC1, BIC2, and HEK293 (Figure S3) (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Author

Shuang Yang – Center for Clinical Mass Spectrometry, Department of Pharmaceutical Analysis, Soochow University, Suzhou, Jiangsu 215123, China; [orcid.org/0000-0001-7958-0594](https://orcid.org/0000-0001-7958-0594); Phone: +1 (410) 320-4681; Email: [yangs2020@suda.edu.cn](mailto:yangs2020@suda.edu.cn)

### Authors

Yan Wang – Mass Spectrometry Facility, National Institute of Dental and Craniofacial Research, National Institutes of Health, Bethesda, Maryland 20892, United States

Zhen Wu – State Key Laboratory of Genetic Engineering, Department of Biochemistry, School of Life Sciences, Fudan University, Shanghai 200438, China

Wenhua Hu – Center for Clinical Mass Spectrometry, Department of Pharmaceutical Analysis, Soochow University, Suzhou, Jiangsu 215123, China

Piliang Hao – School of Life Science and Technology, ShanghaiTech University, Shanghai 201210, China; [orcid.org/0000-0002-3632-1573](https://orcid.org/0000-0002-3632-1573)

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acsomega.1c01785>

### Author Contributions

S.Y. conducted the experiments, analyzed data, and wrote the paper. S.Y. and Y.W. designed the experiments. Y.W., Z.W., W.W., and P.H. conducted LC-MS and MALDI-MS. Y.W. conducted part of data analysis using Byos. S.Y. revised the paper.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

This work was supported by the Intramural Research Program of the NIDCR, National Institutes of Health (ZIA-000751-02 to Y.W. and 1-ZIA-DE000739-05 to L.A.T.) and Soochow University Start-up funding.

## ■ REFERENCES

(1) Li, R.; Pei, S.; Chen, B.; Song, Y.; Zhang, T.; Yang, W.; Shaman, J. Substantial undocumented infection facilitates the rapid dissem-



ination of novel coronavirus (SARS-CoV-2). *Science* **2020**, *368*, 489–493.

(2) 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.

(3) Li, F.; Li, W.; Farzan, M.; Harrison, S. C. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. *Science* **2005**, *309*, 1864–1868.

(4) Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Krüger, N.; Herrler, T.; Erichsen, S.; Schiergens, T. S.; Herrler, G.; Wu, N.-H.; Nitsche, A.; Müller, M. A.; Drosten, C.; Pöhlmann, S. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* **2020**, *181*, 271–280.

(5) Wan, Y.; Shang, J.; Graham, R.; Baric, R. S.; Li, F. Receptor recognition by the novel coronavirus from Wuhan: an analysis based on decade-long structural studies of SARS coronavirus. *J. Virol.* **2020**, *94*, No. e00127-20.

(6) Heurich, A.; Hofmann-Winkler, H.; Gierer, S.; Liepold, T.; Jahn, O.; Pöhlmann, S. TMPRSS2 and ADAM17 cleave ACE2 differentially and only proteolysis by TMPRSS2 augments entry driven by the severe acute respiratory syndrome coronavirus spike protein. *J. Virol.* **2014**, *88*, 1293–1307.

(7) Zhang, H.; Rostami, M. R.; Leopold, P. L.; Mezey, J. G.; O’Beirne, S. L.; Strulovici-Barel, Y.; Crystal, R. G. Expression of the SARS-CoV-2 ACE2 receptor in the human airway epithelium. *Am. J. Respir. Crit. Care Med.* **2020**, *202*, 219–229.

(8) Kawase, M.; Shirato, K.; van der Hoek, L.; Taguchi, F.; Matsuyama, S. Simultaneous treatment of human bronchial epithelial cells with serine and cysteine protease inhibitors prevents severe acute respiratory syndrome coronavirus entry. *J. Virol.* **2012**, *86*, 6537–6545.

(9) Lukassen, S.; Chua, R. L.; Trefzer, T.; Kahn, N. C.; Schneider, M. A.; Muley, T.; Winter, H.; Meister, M.; Veith, C.; Boots, A. W.; Hennig, B. P.; Kreuter, M.; Conrad, C.; Eils, R. SARS-CoV-2 receptor ACE 2 and TMPRSS 2 are primarily expressed in bronchial transient secretory cells. *EMBO J.* **2020**, *39*, No. e105114.

(10) Sungnak, W.; Huang, N.; Bécavin, C.; Berg, M.; Queen, R.; Litvinukova, M.; Talavera-López, C.; Maatz, H.; Reichart, D.; Sampaziotis, F.; Worlock, K. B.; Yoshida, M.; Barnes, J. L. SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes. *Nat. Med.* **2020**, *26*, 681–687.

(11) Zhou, L.; Xu, Z.; Castiglione, G. M.; Soiberman, U. S.; Eberhart, C. G.; Duh, E. J. ACE2 and TMPRSS2 are expressed on the human ocular surface, suggesting susceptibility to SARS-CoV-2 infection. *Ocul. Surf.* **2020**, *18*, 537–544.

(12) Zang, R.; Castro, M. F. G.; McCune, B. T.; Zeng, Q.; Rothlauf, P. W.; Sonnek, N. M.; Liu, Z.; Brulois, K. F.; Wang, X.; Greenberg, H. B.; Diamond, M. S.; Ciorba, M. A.; Whelan, S. P. J.; Ding, S. TMPRSS2 and TMPRSS4 promote SARS-CoV-2 infection of human small intestinal enterocytes. *Sci. Immunol.* **2020**, *5*, No. eabc3582.

(13) Hasan, A.; Paray, B. A.; Hussain, A.; Qadir, F. A.; Attar, F.; Aziz, F. M.; Sharifi, M.; Derakhshankhah, H.; Rasti, B.; Mehrabi, M.; Shahpasand, K.; Saboury, A. A.; Falahati, M. A review on the cleavage priming of the spike protein on coronavirus by angiotensin-converting enzyme-2 and furin. *J. Biomol. Struct. Dyn.* **2021**, *39*, 3025–3033.

(14) Belouzard, S.; Millet, J. K.; Licitra, B. N.; Whittaker, G. R. Mechanisms of coronavirus cell entry mediated by the viral spike protein. *Viruses* **2012**, *4*, 1011–1033.

(15) Ibrahim, I. M.; Abdelmalek, D. H.; Elshahat, M. E.; Elfiky, A. A. COVID-19 spike-host cell receptor GRP78 binding site prediction. *J. Infect.* **2020**, *80*, 554–562.

(16) Wang, Q.; Zhang, Y.; Wu, L.; Niu, S.; Song, C.; Zhang, Z.; Lu, G.; Qiao, C.; Hu, Y.; Yuen, K.-Y.; Wang, Q.; Zhou, H.; Yan, J.; Qi, J. Structural and functional basis of SARS-CoV-2 entry by using human ACE2. *Cell* **2020**, *181*, 894–904.

(17) Wu, A.; Peng, Y.; Huang, B.; Ding, X.; Wang, X.; Niu, P.; Meng, J.; Zhu, Z.; Zhang, Z.; Wang, J.; Sheng, J.; Quan, L.; Xia, Z.; Tan, W.; Cheng, G.; Jiang, T. Genome composition and divergence of the

novel coronavirus (2019-nCoV) originating in China. *Cell Host Microbe* **2020**, *27*, 325–328.

(18) 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.

(19) 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. Vulnerabilities in coronavirus glycan shields despite extensive glycosylation. *Nat. Commun.* **2020**, *11*, No. 2688.

(20) 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.

(21) Sanda, M.; Morrison, L.; Goldman, R. N and O glycosylation of the SARS-CoV-2 spike protein. *Anal. Chem.* **2021**, *93*, 2003–2009.

(22) Bagdonaite, I.; Thompson, A. J.; Wang, X.; Søgaard, M.; Fougereux, C.; Frank, M.; Diedrich, J. K.; Yates, J. R.; Salanti, A.; Vakhrushev, S. Y.; Paulson, J. C.; Wandall, H. H. Site-specific O-glycosylation analysis of SARS-CoV-2 spike protein produced in insect and human cells. *Viruses* **2021**, *13*, No. 551.

(23) An, Y.; Ringer, J. A.; Jarvis, D. L.; Jing, X.; Ye, Z.; Aumiller, J. J.; Eichelberger, M.; Cipollo, J. F. Comparative glycomics analysis of influenza hemagglutinin (H5N1) produced in vaccine relevant cell platforms. *J. Proteome Res.* **2013**, *12*, 3707–3720.

(24) Sun, X.; Jayaraman, A.; Mani Prasad, P.; Raman, R.; Houser, K. V.; Pappas, C.; Zeng, H.; Sasisekharan, R.; Katz, J. M.; Tumpey, T. M. N-linked glycosylation of the hemagglutinin protein influences virulence and antigenicity of the 1918 pandemic and seasonal H1N1 influenza A viruses. *J. Virol.* **2013**, *87*, 8756–8766.

(25) Amanat, F.; Krammer, F. SARS-CoV-2 vaccines: status report. *Immunity* **2020**, *52*, 583–589.

(26) Palmberger, D.; Wilson, I. B. H.; Berger, I.; Grabherr, R.; Rendic, D. SweetBac: a new approach for the production of mammalianised glycoproteins in insect cells. *PLoS One* **2012**, *7*, No. e34226.

(27) Li, W.; De Schutter, K.; Van Damme, E. J. M.; Smaghe, G. Synthesis and biological roles of O-glycans in insects. *Glycoconjugate J.* **2020**, *37*, 47–56.

(28) Wang, H.; Li, X.; Li, T.; Zhang, S.; Wang, L.; Wu, X.; Liu, J. The genetic sequence, origin, and diagnosis of SARS-CoV-2. *Eur. J. Clin. Microbiol. Infect. Dis.* **2020**, *39*, 1629–1635.

(29) Chen, Y.; Rajashankar, K. R.; Yang, Y.; Agnihothram, S. S.; Liu, C.; Lin, Y.-L.; Baric, R. S.; Li, F. Crystal structure of the receptor-binding domain from newly emerged Middle East respiratory syndrome coronavirus. *J. Virol.* **2013**, *87*, 10777–10783.

(30) Yang, S.; Jankowska, E.; Kosikova, M.; Xie, H.; Cipollo, J. Solid-phase chemical modification for sialic acid linkage analysis: application to glycoproteins of host cells used in influenza virus propagation. *Anal. Chem.* **2017**, *89*, 9508–9517.

(31) 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.

(32) Jarvis, D. L. Developing baculovirus-insect cell expression systems for humanized recombinant glycoprotein production. *Virology* **2003**, *310*, 1–7.

(33) Watanabe, S.; Kokuho, T.; Takahashi, H.; Takahashi, M.; Kubota, T.; Inumaru, S. Sialylation of N-glycans on the recombinant proteins expressed by a baculovirus-insect cell system under  $\beta$ -N-acetylglucosaminidase inhibition. *J. Biol. Chem.* **2002**, *277*, 5090–5093.

(34) Kong, Y.; Joshi, H. J.; Schjoldager, K. T.-B. G.; Madsen, T. D.; Gerken, T. A.; Vester-Christensen, M. B.; Wandall, H. H.; Bennett, E. P.; Levery, S. B.; Vakhrushev, S. Y.; Clausen, H. Probing polypeptide GalNAc-transferase isoform substrate specificities by in vitro analysis. *Glycobiology* **2015**, *25*, 55–65.

(35) Mohl, J. E.; Gerken, T. A.; Leung, M.-Y. ISOGLyP: de novo prediction of isoform specific mucin-type O-glycosylation. *Glycobiology* **2021**, *31*, 168–172.

- (36) Hema Thanka Christlet, T.; Veluraja, K. Database analysis of O-glycosylation sites in proteins. *Biophys. J.* **2001**, *80*, 952–960.
- (37) Bennett, E. P.; Mandel, U.; Clausen, H.; Gerken, T. A.; Fritz, T. A.; Tabak, L. A. Control of mucin-type O-glycosylation: a classification of the polypeptide GalNAc-transferase gene family. *Glycobiology* **2012**, *22*, 736–756.
- (38) May, C.; Ji, S.; Syed, Z. A.; Revoredo, L.; Daniel, E. J. P.; Gerken, T. A.; Tabak, L. A.; Samara, N. L.; Ten Hagen, K. G. Differential splicing of the lectin domain of an O-glycosyltransferase modulates both peptide and glycopeptide preferences. *J. Biol. Chem.* **2020**, *295*, 12525–12536.
- (39) Pratt, M. R.; Hang, H. C.; Ten Hagen, K. G.; Rarick, J.; Gerken, T. A.; Tabak, L. A.; Bertozzi, C. R. Deconvoluting the functions of polypeptide N- $\alpha$ -acetylgalactosaminyltransferase family members by glycopeptide substrate profiling. *Chem. Biol.* **2004**, *11*, 1009–1016.
- (40) Moremen, K. W.; Ramiah, A.; Stuart, M.; Steel, J.; Meng, L.; Forouhar, F.; Moniz, H. A.; Gahlay, G.; Gao, Z.; Chapla, D.; et al. Human glycosylation enzymes for enzymatic, structural and functional studies. *Nat. Chem. Biol.* **2018**, *14*, 156–162.
- (41) Marchal, I.; Jarvis, D. L.; Cacan, R.; Verbert, A. Glycoproteins from insect cells: sialylated or not? *Biol. Chem.* **2001**, *382*, 151–159.
- (42) Kubelka, V.; Altmann, F.; Kornfeld, G.; Marz, L. Structures of the N-linked oligosaccharides of the membrane glycoproteins from three lepidopteran cell lines (Sf-21, IZD-Mb-0503, Bm-N). *Arch. Biochem. Biophys.* **1994**, *308*, 148–157.
- (43) Stencel-Baerenwald, J. E.; Reiss, K.; Reiter, D. M.; Stehle, T.; Dermody, T. S. The sweet spot: defining virus–sialic acid interactions. *Nat. Rev. Microbiol.* **2014**, *12*, 739–749.
- (44) Schwegmann-Weßels, C.; Herrler, G. Sialic acids as receptor determinants for coronaviruses. *Glycoconjugate J.* **2006**, *23*, 51–58.
- (45) Lan, J.; Ge, J.; Yu, J.; Shan, S.; Zhou, H.; Fan, S.; Zhang, Q.; Shi, X.; Wang, Q.; Zhang, L.; Wang, X. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **2020**, *581*, 215–220.
- (46) Shang, J.; Ye, G.; Shi, K.; Wan, Y.; Luo, C.; Aihara, H.; Geng, Q.; Auerbach, A.; Li, F. Structural basis of receptor recognition by SARS-CoV-2. *Nature* **2020**, *581*, 221–224.
- (47) Li, F. Structural analysis of major species barriers between humans and palm civets for severe acute respiratory syndrome coronavirus infections. *J. Virol.* **2008**, *82*, 6984–6991.
- (48) 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.
- (49) Han, D. P.; Lohani, M.; Cho, M. W. Specific asparagine-linked glycosylation sites are critical for DC-SIGN and L-SIGN-mediated severe acute respiratory syndrome coronavirus entry. *J. Virol.* **2007**, *81*, 12029–12039.
- (50) Bernardi, A.; Huang, Y.; Harris, B.; Xiong, Y.; Nandi, S.; McDonald, K. A.; Faller, R. Development and simulation of fully glycosylated molecular models of ACE2-Fc fusion proteins and their interaction with the SARS-CoV-2 spike protein binding domain. *PLoS One* **2020**, *15*, No. e0237295.
- (51) Mishin, V. P.; Novikov, D.; Hayden, F. G.; Gubareva, L. V. Effect of hemagglutinin glycosylation on influenza virus susceptibility to neuraminidase inhibitors. *J. Virol.* **2005**, *79*, 12416–12424.
- (52) Shen, R.; Raska, M.; Bimczok, D.; Novak, J.; Smith, P. D. HIV-1 envelope glycan moieties modulate HIV-1 transmission. *J. Virol.* **2014**, *88*, 14258–14267.
- (53) Goh, J. B.; Ng, S. K. Impact of host cell line choice on glycan profile. *Crit. Rev. Biotechnol.* **2018**, *38*, 851–867.
- (54) Le Ru, A.; Jacob, D.; Transfiguracion, J.; Ansonge, S.; Henry, O.; Kamen, A. A. Scalable production of influenza virus in HEK-293 cells for efficient vaccine manufacturing. *Vaccine* **2010**, *28*, 3661–3671.
- (55) Tian, J.-H.; Patel, N.; Haupt, R.; Zhou, H.; Weston, S.; Hammond, H.; Logue, J.; Portnoff, A. D.; Norton, J.; Guebre-Xabier, M.; Zhou, B.; Jacobson, K.; Maciejewski, S.; Khatoon, R.; Wisniewska, M.; Moffitt, W.; Kluepfel-Stahl, S.; Ekechukwu, B.; Papin, J.; Boddapati, S.; Jason Wong, C.; Piedra, P. A.; Frieman, M. B.; Massare, M. J.; Fries, L.; Bengtsson, K. L.; Stertman, L.; Ellingsworth, L.; Glenn, G.; Smith, G. SARS-CoV-2 spike glycoprotein vaccine candidate NVX-CoV2373 immunogenicity in baboons and protection in mice. *Nat. Commun.* **2021**, *12*, No. 372.
- (56) Harrison, R. L.; Jarvis, D. L. Protein N-glycosylation in the baculovirus–insect cell expression system and engineering of insect cells to produce “mammalianized” recombinant glycoproteins. *Adv. Virol. Res.* **2006**, *68*, 159–191.
- (57) Wang, Q.; Stuczynski, M.; Gao, Y.; Betenbaugh, M. J. Strategies for Engineering Protein N-Glycosylation Pathways in Mammalian Cells. In *Methods in Molecular Biology*; Springer, 2015; Vol. 1321, pp 287–305.
- (58) Tai, W.; He, L.; Zhang, X.; Pu, J.; Voronin, D.; Jiang, S.; Zhou, Y.; Du, L. Characterization of the receptor-binding domain (RBD) of 2019 novel coronavirus: implication for development of RBD protein as a viral attachment inhibitor and vaccine. *Cell. Mol. Immunol.* **2020**, *17*, 613–620.
- (59) Yang, S.; Wu, W. W.; Shen, R.-F.; Bern, M.; Cipollo, J. Identification of sialic acid linkages on intact glycopeptides via differential chemical modification using IntactGIG-HILIC. *J. Am. Soc. Mass Spectrom.* **2018**, *29*, 1273–1283.
- (60) Yang, S.; Li, Y.; Shah, P.; Zhang, H. Glycomic analysis using glycoprotein immobilization for glycan extraction. *Anal. Chem.* **2013**, *85*, 5555–5561.
- (61) Ceroni, A.; Maass, K.; Geyer, H.; Geyer, R.; Dell, A.; Haslam, S. M. GlycoWorkbench: a tool for the computer-assisted annotation of mass spectra of glycans. *J. Proteome Res.* **2008**, *7*, 1650–1659.