Mass spectrometry analysis of newly emerging coronavirus HCoV-19 spike S protein and human ACE2 reveals camouflaging glycans and unique post-translational modifications

Zeyu Sun[1], Keyi Ren[1], Xing Zhang[2], Jinghua Chen[2], Zhengyi Jiang[1], Jing Jiang[1], Feiyang Ji[1], Xiaoxi Ouyang[1], Lanjuan Li[1*]

[1] State Key Laboratory for Diagnosis and Treatment of Infectious Diseases, National Clinical Research Center for Infectious Diseases, Collaborative Innovation Center for Diagnosis and Treatment of Infectious Disease, The First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, Zhejiang, China.

[2] Department of Biophysics & Center of Cryo-Electron Microscopy, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China

Email:

Zeyu Sun, zeyusun@zju.edu.cn
Keyi Ren, 13735509115@163.com
Xing Zhang, xzhang1999@zju.edu.cn
Chen, cjh188@zju.edu.cn
Zhengyi Jiang, 11518148@zju.edu.cn
Jing Jiang, jiangjing_315@126.com
Feiyang Ji, feiyangji@zju.edu.cn
Xiaoxi Ouyang, ouyangxiaoxidd@163.com
Lanjuan Li, ljli@zju.edu.cn

*Correspondence: Lanjuan Li, The First Affiliated Hospital, College of Medicine, Zhejiang University, No.79 Qingchun Road, Shangcheng District, 310003 Hangzhou,

Zhejiang Province, China

E-mail: ljli@zju.edu.cn

Telephone: +86-571-87236458

Fax: +86-571-87236459.

Abstract

The pneumonia-causing COVID-19 pandemia has prompt worldwide efforts to understand its biological and clinical traits of newly identified HCoV-19 virus. In this study, post-translational modification (PTM) of recombinant HCoV-19 S and hACE2 were characterized by LC-MSMS. We revealed that both proteins were highly decorated with specific proportions of N-glycan subtypes. Out of 21 possible glycosites in HCoV-19 S protein, 20 were confirmed completely occupied by N-glycans, with oligomannose glycans being the most abundant type. All 7 possible glycosylation sites in hACE2 were completely occupied mainly by complex type N-glycans. However, we showed that glycosylation did not directly contribute to the binding affinity between SARS-CoV spike protein and hACE2. Additionally, we also identified multiple sites methylated in both proteins, and multiple prolines in hACE2 were converted to hydroxylproline. Refined structural models were built by adding N-glycan and PTMs to recently published cryo-EM structure of the HCoV-19 S and hACE2 generated with glycosylation sites in the vicinity of binding surface. The PTM and glycan maps of both HCoV-19 S and hACE2 provide additional structural details to study mechanisms underlying host attachment, immune response mediated by S protein and hACE2, as well as knowledge to develop remedies and vaccines desperately needed nowadays.

Keywords: N-glycosylation, COVID-19, Spike protein, hACE2, Structure

Introduction

The on-going coronavirus pandemia since December 2019 (COVID-19) has now become a global health emergence (Dong et al., 2020; Wu et al., 2020b; Zhou et al., 2020b). The disease is caused by a newly identified beta-coronavirus HCoV-19 (also known as SARS-CoV-2) which is closely related to severe acute respiratory syndrome coronavirus (SARS-CoV). Like the SARS, the HCoV-19 causes lower respiratory tract infection (LRTI) which may eventually develop to atypical pneumonia that requires intensive care and mechanical ventilation (Guan et al., 2020; Wang et al., 2020; Zhou et al., 2020a). Compared to SARS-CoV, the HCoV-19 is more contagious and is infecting more population worldwide thus causing substantially more casualties (Guan et al., 2020; Zhou et al., 2020a). Although possible viral reservoirs including bats and pangolin have been suggested for HCoV-19 (Zhang et al., 2020; Zhou et al., 2020b), more investigations are need to ascertain its source and routes of zoonotic transmission (Wu et al., 2020a; Xu, 2020; Zhang and Holmes, 2020). To date, no approved vaccine or remedy specific for coronaviruses infection, including that of HCoV-19, is available.

Akin to the SARS-CoV, the virion surface spike glycoprotein encoded by HCoV-19 S gene is essential for target cell attachment and fusion processes (Hoffmann et al., 2020; Walls et al., 2020a; Wan et al., 2020). Given the importance of S protein in COVID-19 pathogenesis and its potential immunogenicity for vaccine development, global efforts have been made to elucidate the structure of S protein shortly after the publishing of the first HCoV-19 sequence (Walls et al., 2020a; Wrapp et al., 2020; Wu et al., 2020b). The HCoV-19 S glycoprotein is a 1273-amino acid precursor polypeptide and can be cleaved by host cell proteases (cathepsin L, TMPRSS2) into S1 fragment that contains the receptor binding domain (RBD) to attach host receptor human Angiotensin I Converting Enzyme 2 (hACE2), and the S2 fragment responsible for the subsequent membrane fusion (Hoffmann et al., 2020; Walls et al., 2020a; Wu et al., 2020b). The S protein is predicted to have a cleavable N-terminal signal sequence (1-15), which presumably directs the protein toward the endoplasmic reticulum (ER) for extensive glycan decoration before virion packing.

Glycosylation is one of the most prominent post-translational modifications (PTM) in many viral spike or envelop proteins, and has been shown to mediate host attachment, immune response as well as virion packaging and budding (Chang and Zaia, 2019; de Groot, 2006; Fukushi et al., 2012; Li et al., 2017; Parsons et al., 2019; Raman et al., 2016; Shih et al., 2006; York et al., 2019; Zheng et al., 2018; Zhou et al., 2010). In particular, the binding of coronavirus S proteins to their respective receptors has been shown to be mediated by its oligomannose N-glycan (Li et al., 2017; Parsons et al., 2019; Zheng et al., 2018). Additionally, C-type lectin DC-SIGN and L-SIGN can enhance viral entry via their binding to the S protein glycans (Han et al., 2007; Jeffers et al., 2004). Therefore, numerous antiviral strategies have been designed to interfere protein glycosylation or glycan-based interaction (Shih et al., 2006; Vincent et al., 2005; Zheng et al., 2020; Zhou et al., 2010), and glycosylation is considered as a key aspect to develop effective vaccine (Chen et al., 2014; Kumar et al., 2020).

Like the SARS-CoV spike protein which contains 23 N-linked glycosylation sequons (N-X-S/T, X≠P), HCoV-19 spike protein is predicted to host 22 per protomer or 66 per trimer (Walls et al., 2020a). However, the potential glycosites pattern in S1 of HCoV-19 S protein is different to that of SARS-CoV, while the glycosylation sites in S2 region are significant conserved between HCoV-19 and SARS-CoV. The difference of S1 glycosylation pattern may link to biological and clinical characteristics of HCoV-19 which is profoundly different to other coronavirus. In this study, we report a comprehensive N-glycosylation profile, as well as other PTMs, of HCoV-19 S protein and hACE2 elucidated by high resolution mass spectrometry analyses, based on which reinterpretation of current HCoV-19 S protein structural model was made to highlight important glycan features related to COVID-19 pathogenesis. Nonetheless, we demonstrate that the binding of HCoV-19 S protein and hACE2 is independent on their glycosylation status.

Methods

Expression and purification of HCoV-19 spike ectodomain and hACE2

HCoV-19 S gene (virus isolate: Wuhan Hu-1; GenBank number QHD43416.1) was synthesized (Genscript) with codons optimized for insect cell expression. Its ectodomain (Val16-Pro1213) was cloned into pFastBac vector (Life Technologies Inc.) with a N-terminal honeybee melittin signal peptide and C-terminal His6 and Flag tags. HCoV-19 S protein was expressed in Sf9 insect cells using the Bac-to-Bac system (Life Technologies Inc.) and harvested from cell culture medium followed by purification procedure using Ni-NTA column and Superdex 200 gel filtration column (GE Healthcare) in tandem. The extracellular domain of hACE2 (Gln18-Ser740, NP_068576.1) with C-terminal Fc tag was expressed in HEK 293 cells were purified by protein A sepharose beads (GE Healthcare).

SDS-PAGE analysis

To test their glycosylation status of HCoV-19 spike and hACE2 protein, both proteins were deglycosylated by PNGase F (NEB, 1:50) overnight at 37°C in PBS. Both the glycosylated and deglycosylated forms of HCoV-19 spike and hACE2 protein were then analyzed by 15% polyacrylamide gel electrophoresis (SDS-PAGE) followed by Coomassie blue staining.

Glycopeptide sample preparation

Proteins were first digested into tryptic peptides according to (Sun et al., 2016). Briefly, hACE2 and S protein were first deduced by 10 mM dithiothreitol in 50 mM ammonium bicarbonate (ABC) for 45 min at room temperature and then alkylated by iodoacetamide for 45 mins at room temperature in the dark. Proteins were then cleaned up by acetone precipitation and resuspended in 50 mM ABC. The alkylated glycoproteins were then digested for 14 h at 37°C using sequencing-grade trypsin, chymotrypsin or endoproteinase Lys-C, all purchased from Promega, with a protein:protease ratio of 1:50 in 50 mM ABC. The peptides were desalted using HLB columns (Waters).

The peptide samples were separated into duplicates. The first duplicate was deglycosylated by PNGase F (NEB, 1:100) overnight at 37°C in 50 mM ABC prepared in pure H_2O^{18}. The second duplicate was used to enrich intact N-glycopeptides by hydrophilic interaction liquid chromatography (HILIC). Briefly, the GlycoWorks (Waters) cartridges were pre-conditioned with loading buffer comprised of 15 mM ammonium acetate (AmA) and 0.1% trifluoroacetic acid (TFA) in 80% Acetonitrile (ACN). Peptides in loading buffer were applied to the cartridges and unbound flow-through (FT) fraction was collected. The column was washed two times with loading buffer. Glycopeptides were eluted in 0.1% TFA. All peptide fractions were desalted by house-made C18 stagetips before LC-MSMS analyses.

LC-MSMS experiment

About 500 ng of peptides were analysis using an Ultimate 3000 nanoflow liquid chromatography system (Thermo Scientific, USA) connected to a hybrid Q-Exactive HFX mass spectrometry (Thermo Scientific, USA). The mass spectrometer was operated in data-dependent mode with a full scan MS spectra followed by MS2 scans recording the top 20 most intense precursors sequentially isolated for fragmentation using high energy collision dissociation (HCD). The MS and MS/MS spectra were recorded using the Xcalibur software 2.3 (Thermo Scientific, USA). Detail parameters for LC separation and mass spectrometry acquisition can be found in Supplementary Table 1 and 2, respectively.

Bioinformatics

The acquired MS raw files from deglycosylated peptides were searched by MaxQuant (version 1.6.10.43) against the human ACE2 sequence from UniProtKB and HCoV-19 S sequence (YP_009724390.1_3) from NCBI. Cysteine carbamidomethylation was set as fixed modification, while methionine oxidation and O^{18} deamidation on glutamic acid were set as variable modification. Trypsin with up to two missed cleavages was set. Mass tolerance of 15 ppm and 4.5 ppm were set for

first and main search, respectively. For comprehensive PTM analysis, phosphorylation (S/T/Y), acetylation (K), methylation (K/R/E), succinylation (K), crotonylation (K), farnesylation (K/Nterm), myristoylation (K/Nterm), palmitoylation or prenylation, glycosylphosphatidylinositol and oxidation on proline were individually investigated in parallel database searches for respective variable modification. Peptide level 1% FDR was set to filter the result. Confident identification of PTM was based on localization probability of 99%. Site occupancy for each PTM was calculated by dividing the peak intensities of the modified peptides (MP) and corresponding non-modified peptides (NP) using the equation: MP/(MP+NP)*100.

To investigate N-glycosylation forms, acquired MS raw files from HILIC experiments were searched by pGlyco (version 2.2.2) (Liu et al., 2017) against the same hACE2 and HCoV-19 S sequences. Cysteine carbamidomethylation was set as fixed modification, while methionine oxidation was set as variable modification. Trypsin with up to two missed cleavages was set. Mass tolerance of 5 ppm and 15 ppm were set for precursor and fragment mass tolerance, respectively. Potential glycan fragments within MS2 spectra were annotated by built-in pGlyco.gdb glycan structure database (Liu et al., 2017). Precursor intensity of each glycopeptide was extracted by MaxQuant feature detection algorithm.

Structural model refinement

Glycan structure in pdb format was downloaded or predicted by GLYCAM-Web (https://dev.glycam.org/gp). The N-linked glycan models of hACE2 and trimeric spike protein were made by manually adding the MS-identified glycan at each site based on previous models (PDB codes 6M18 for hACE2 and 6VXX for spike protein) within Coot (Emsley et al., 2010). The modified residues within each model were generated with Coot (Emsley et al., 2010). In addition, model of HCoV-19 S protein and hACE2 complex (PDB 6M0J) were used to shown location of PTM surrounding the binding area. All structural figures were performed by PyMOL (Schrödinger Inc. U.S.A.).

200

201 **Evaluation of binding of HCoV-19 S protein and hACE2 by bio-layer**
202 **interferometry (BLI)**

203 Binding of HCoV-19 S protein and hACE2 was measured on an Octet
204 Re96E (ForteBio) interferometry system. Briefly, HCoV-19 S protein were
205 immobilized using aminopropylsilan biosensors (18-5045, ForteBio). To evaluate the
206 influence of glycosylation of HCoV-19 S on binding, S protein were first incubated
207 with either 1000U/mL PNGase F (NEB) or deactivated PNGase F at 37 °C for 14 h.
208 After washing by kinetic buffer (1x PBS, pH 7.4, 0.01% BSA, and 0.002% Tween 20),
209 the association between HCoV-19 S and hACE2 was measured for 180 seconds at
210 30°C by exposing sensors to hACE2 in kinetic buffer at concentration of 12.5, 25, 50,
211 100, 200 nM, respectively. After binding phase, the sensors were exposed to 1x
212 kinetic buffer for dissociation for 300 seconds at 30°C. Signal baseline was subtracted
213 before data fitting using the equimolar binding model. Mean k_{on}, k_{off} values were
214 determined with a global fit model using all data. A parallel experiment was also
215 performed using hACE2 loaded sensor incubated with HCoV-19 S protein solution.
216 Analogously, immobilized hACE2 was pretreated by either PNGase F or deactivated
217 PNGase F to evaluate the influence of hACE2 glycosylation on binding. In all
218 experiment, PNGase F was deactivated by heating at 75 °C for 10 min.

219

Results

Determination of N-linked glycosylation sites on HCoV-19 S protein and hACE2

As shown in (Figure 1A), PNGase F deglycosylation resulted in deceasing molecular weight of both HCoV-19 S protein and hACE2 on SDS-PAGE. Digestion by PNGase F releases N-linked glycans from Asn within the N-X-S/T motif, which also resulted in Asn deamidation by losing one hydrogen and one nitrogen while incorporating one oxygen derived from the solvent (Bailey et al., 2012). To perform a thorough survey of glycosylated sites, protease digested peptides from both proteins were subjected to PNGase F deglycosylation in H_2O^{18} that enables incorporation of O^{18} and leads to a +2.98 Da mass increment, which was used to mark glycosylated sites.

Analysis of the deglycosylated peptides from S protein by LC-MS/MS confirmed 20 N-linked glycosylation sites, including N61, N74, N122, N165, N234, N282, N331, N343, N603, N616, N657, N709, N717, N801, N1074, N1098, N1134, N1158, N1173 and N1194 (Table 1). Two remaining N-X-S/T sites, the N17 and N149 were not identified in any deglycosylated peptide. However, peptide with glycosylated N149 was identified directly without PNGase F treatment as described later, therefore leading to a total of 21 glycosylation sites out of 22 potential sites with N-X-S/T sequons in HCoV-19 S protein (Figure 1B). In addition, quantitative analysis of site occupancy showed there were 18 sites completely glycosylated, while N603, N657 reached 43% and 74% occupancy (Table 1). Example mass spectra showing evidence of N331 and N343 glycosylation was shown in (Figure 1C).

All 7 possible glycosylation sites in hACE2, N53, N90, N103, N322, N432, N546 and N690 were confirmed as glycosylated in our experiments (Table 1). As example, mass spectra of N90 glycosylation were shown in Fig 1D. Quantitative analysis of site occupancy showed all 7 sites were completely glycosylated (Table 1). All these data suggested that both HCoV-19 S protein and hACE2 are highly decorated by N-glycans.

Identification of PTMs of HCoV-19 S protein and hACE2

To unveil possible post-translational modification (PTM) other than glycosylation, LC-MSMS data of deglycosylated peptides from both HCoV-19 S protein and hACE2 were subjected to several rounds of spectra-database matching for common PTMs. The major PTM found in both proteins was methylation on lysine, arginine or glutamic acid, as summarized in Table 1. Using occupancy >50% as criteria to identify site dominantly decorated with PTMs, we found 78R, 224E, 654E and 661E in HCoV-19 S protein, 57E, 68K and 329E in hACE2 is highly methylated. In addition, proline at site 253, 263, 321 and 346 in hACE2 can be found prominently oxidized and converted to hydroxyproline. However, common PTMs such as phosphorylation, acetylation or other acylations were not found in HCoV-19 S protein and hACE2.

Global N-glycosylation profile of HCoV-19 S protein and hACE2

To resolve glycan camouflage on surface of HCoV-19 S protein and hACE2, intact glycopeptides derived from protease digestion and fractionated by HILIC SPE were directly subject to LC-MSMS analysis specifically designed to detect peptides with extra molecular weight due to N-glycan attachment. Due to the highly heterogeneous nature of glycan chain, in our study a unique glycopeptides was designated as combination of both unique peptide sequence with a specific N-glycan composition. According to this criteria, 419 and 467 unique N-glycopeptides were identified from HCoV-19 S proteins and hACE2, respectively (Table 2 and 3). Out of 20 HCoV-19 S glycosylated sides identified by PNGase F experiment, 19 were confirmed by intact glycopeptide profiling. Despite all 7 glycosylation sites in hACE2 were identified previously in deglycosylated peptides labeled with deamidation, N-glycan profile was obtained in 5 sites: N90, N103, N432, N546 and N690. Example N-glycopeptides spectra from HCoV-19 S and hACE2 could be found in Supplementary Figure 1.

A total of 144 N-glycans were found in HCoV-19 S protein, with majority of them containing the common N-acetylglucosamine core (Table 2). All N-glycosites in S protein were attached with multiple types of N-glycans, with N343 decorated by the

most diverse N-glycans. The HCoV-19 S N-glycan composition preferentially comprising pauci- or high-mannose type oligosaccharides, except in four N-glycosites (N73, N343, N717 and N1173) containing a high proportion of complex and hybrid N-glycans (Figure 3A, 3B and Table 2). By LC-MS intensity, pauci-mannose Hex3HexNAc2 was the most common N-glycan in HCoV-19 S protein. Interestingly, there was no evidence of sialic acid component in the N-glycan of HCoV-19 S protein.

A total of 220 N-glycans was found in hACE2, with 78.2% of them being complex type (Figure 3A), while hybrid, high-mannose and pauci-mannose glycan only constitute 10.9%, 5.0% and 5.9%, respectively (Figure 3B and Table 3). By LC-MS intensity, the most dominant hACE2 glycan was NeuGc1Hex5HexNAc4 which was predicted to be a bi-antennary complex containing a sialic acid end. N-glycan profile in all hACE2 sites was primarily complex type. It is notable that N90, N103 and N690 all contain more than 100 types of N-glycans, with their most dominant N-glycan form all containing sialic acid.

Collectively, our LC-MSMS data confirms that both HCoV-19 spike protein and its receptor hACE2 are indeed heavily N-glycosylated at most of its predicted N-X-S/T sequon. Summary of most dominant N-glycan composition and predicted structure in HCoV-19 spike protein and hACE2 were illustrated in Figure 3C.

Structure modeling of HCoV-19 S-hACE2 complex refined with glycan and PTM details

By adding chemical structures of the most abundant N-glycans in each site based on LC-MSMS results to the most updated cyro-EM model of HCoV-19 s protein and hACE2, we generated atomic models that represented the most likely spatial distribution of the N-glycans on both proteins (Fig. 3A and 3B). Despite 4 sites (N74, N149, N1158, and N1194) of spike protein were not shown in the model for the missing residues in model 6VXX, our model suggested the camouflaging N-glycans shielded more than 2/3 of the HCoV-19 S protein surface, which could potentially lead to host attachment and immune evasion. Additionally, both HCoV-19 s protein

and hACE2 model showed the glycans at N331, N343 of HCoV-19 s protein and N90 of hACE2 were in the proximity of, albeit not exactly inside of, the binding area of both proteins (Fig. 3C and 3D). Model of S-hACE2 complex also showed that 3 methylation sites in hACE2 (57E, 68K and 329E) formed a trident structure which enclosed the contact area formed between K353-R357 of hACE2 and N501 of HCoV-19 S (Supplementary Figure 2).

Binding of HCoV-19 S protein and hACE2 does not depends on N-glycosylation

To understand the contribution of protein glycosylation to the interaction HCoV-19 S protein with hACE2, we compared the binding kinetics and affinity of the purified hACE2 ectodomain to glycosylated and deglycosylated HCoV-19 S-ECD immobilized at the surface of biosensors in a BLI experiment. We found that hACE2 bound to HCoV-19 S pretreated with active or inactive PNGase F with comparable equilibrium dissociation constants Kd of 1.7 nM (Figure 4A) and 1.5 nM (Figure 4B), respectively. The affinity determined in this study for hACE2 binding to HCoV-19 S is in line with (Walls et al., 2020a). Comparable affinity was also observed when HCoV-19 S ECD was bound to the immobilized hACE2 (Kd = 16.7 nM, Figure 4C) or deglycosylated hACE2 (Kd = 18.2 nM, Figure 4D). Detailed summary of BLI binding kinetics can be found in Supplementary Table 3.

Discussion

Glycosylation is a ubiquitous and complex PTM that greatly extend the structural and functional diversity for many proteins. However, comprehensive characterization of protein N-glycosylation is technically challenging. In this study, deglycosylated peptides which are easier to be analyzed by LC-MS than their glycosylated precursors, were profiled to confirmed N-glycosylated sites in HCoV-19 S and hACE2. The O^{18} deamidation modification significantly boost the confidence for glycosite identifications as the isotope was incorporated by PNGase F mediated hydrolysis in H_2O^{18}. Intact N-glycopeptide analyses were then performed to further verify the glycosites, and provided detailed information about the compositional and structural

heterogeneity of N-glycans associated with each site. Our investigation showed that all sites except N17 were highly glycosylated in HCoV-19 S protein. According to the N-linked glycan model of spike protein, a large proportion of the protein surface is covered by glycans, which is consistent with that in previous reports (Walls et al., 2020b; Wrapp et al., 2020). When compared spike proteins from HCoV-19 and SARS-CoV, it is noticeable that the majority of differences of glycosylation sites occurs in the distal S1 subunit, therefore leading to significant difference in glycan profile in the outermost canopy of the virus formed by spike trimer clusters. Alteration of glycosites might be the results of host or environmental selective pressure, and could implicate dramatic difference in viral infectivity, pathogenesis and host responses.

The N glycans in S protein are markedly heterogeneous, which greatly extends conformational flexibility and epitope diversity. While a small proportion of complex and hybrid N-glycans were found in S protein, most sites were occupied by oligomannose glycans, which is in line with previous report on SARS-CoV (Krokhin et al., 2003; Song et al., 2004; Ying et al., 2004). Many viral proteins are hallmarked by high level of oligomannose-type glycans, probably as a sign of incomplete glycan maturation due to high glycosite density that results in steric hindrance (Watanabe et al., 2019). Interestingly, N343, the glycosite closet to the binding surface, is decorated by the most diverse N-glycans and primarily by hybrid and complex forms. Moreover, like in case of SARS-CoV and Marburg viral proteins (Feldmann et al., 1994; Ritchie et al., 2010), we found that sialic acid incorporation in SARS-CoV spike protein glycans is negligible. Regarding the ongoing global efforts to generate vaccine, which primarily targets S protein as the candidate antigen, the diverse glycan forms decorating large section of S trimer surface should be considered by vaccine developers as they could drastically modulate the protein immunogenicity. We should also expect that vaccines generated based on S proteoforms with different glycosylation status could render varying efficacy against future infection.

The binding of coronavirus S proteins to their respective receptors has been shown to be mediated by its oligomannose N-glycan (Li et al., 2017; Parsons et al.,

2019; Zheng et al., 2018). SARS-CoV spike protein contains 3 RBD-associated glycosites (N318, N330 and N357), while HCoV-19 only has 2 (N331 and N343) that surrounds the binding pocket. In contrary to our postulation that N-glycan at these two sites might contribute polar interaction to receptor binding, the BLI binding assay suggested deglycosylation did not change the affinity of SARS-CoV spike protein to hACE2. However, this negative BLI binding results do not exclude the possibility that glycosylation could affect other viral entry steps, including protease cleavage, and glycans mediated interactions with DC/L-SIGN which were documented to enhance viral entry in SARS-CoV (Han et al., 2007; Jeffers et al., 2004). Further investigations are required to ascertain the role of glycosylation regarding the infectivity of HCoV-19.

Antigen glycosylation greatly determines host immune responses. One of the prominent consequence of viral envelope or surface proteins is immune evasion by shield off immunogenic epitopes (Raman et al., 2016; Vigerust and Shepherd, 2007; Yang et al., 2020). The complete occupancy of glycans in most glycosite of HCoV-19 S protein suggest the virus is able to invade the host in a stealth fashion. Successful innate immune evasion by the HCoV-19 at the early stage of infection might explain the long asymptomatic incubation period during which transmissible virion are produced (Guan et al., 2020; Wang et al., 2020). Cases of HCoV-19 reinfection were also reported suggesting the virus is capable to escape from antibody-mediated neutralization (Biswas et al., 2020). On the other hand, protein glycosylation can also facilitate immune response by sensor molecules or antibodies specific for glycan recognition. N-glycans are known ligands for galectins, which triggers inflammation events such as cytokine release, immune cell infiltration (Robinson et al., 2019; Wang et al., 2019), of which may contribute to the HCoV-19 pathogenesis and correlated with the disease severity. Moreover, epidemiological studies have shown that ABO polymorphism are linked to different susceptibility to both HCoV-19 and SARS-CoV infections (Cheng et al., 2005; Zhao et al., 2020). Since the ABH carbohydrate epitopes and viral protein glycans are likely to be synthesized by the same ER-resident glycosylation enzymes, anti-A or -B antibodies could partially block the

viral-host interactions (Guillon et al., 2008), thereby with blood type O individuals more resistant to virus infection. It remains to be explored if the glycosylation status of S protein has implications in the inter-individual variation in terms of viral response and clinical outcome within the infected population.

The glycosylation status of hACE2 was also profiled in this study. All 7 glycosites in hACE2 were completely occupied by glycan, including the N90 in the vicinity of binding surface. However, we found the glycan did not directly contribute to the its binding with HCoV-19 S protein. This is in agreement with previous founding that the disruption of ACE2 glycosylation did not affect its binding of SARS-CoV spike protein (Zhao et al., 2015). But deglycosylated ACE2 did compromise the cellular entry and the subsequent production of infectious SARS-CoV virion (Zhao et al., 2015). Therefore, ACE2 glycosylation was still considered as an important target for intervention of coronavirus infection. In fact, the ability of chloroquine to counter HCoV-19 infection was thought to be the results of its inhibition ACE2 glycosylation in addition to its ability to increase endosomal pH level (Liu et al., 2020; Vincent et al., 2005).

Additional PTM forms other than glycosylation were also investigated in this study. Methylation was identified in several sites in both S protein and hACE2. In particular, we found 57E, 68K and 329E in hACE2 that surrounding its binding site with RBD of S protein are completed methylated. Methylation causes loss of charge and increases hydrophobicity in these sites. Four hydroxyproline (253, 263, 321, 346) were identified in the protease domain of hACE2. The extra hydroxyl group is like to increase hydrophilicity of proline in this extracellular region of hACE2. Further study is needed to investigate possible biological role of these PTMs. We did not found phosphorylation and acetylation in our dataset as previous PTM investigation on SARS-CoV also fail to identify these PTMs (Ying et al., 2004). Fatty acid acylations were also been considered in our study since they are commonly found in surface viral protein to facilitate membrane fusion and virus entry. However, our study found no evidence of acylation PTM presence in HCoV-19 S protein. Given that our multi-protease proteomic experiment provides almost full coverage for both proteins,

we believe the PTMome of HCoV-19 S protein and hACE2 mainly consist of glycosylation, methylation and proline oxidation.

Acknowledgments

We thank the proteomics and metabolomics platform in the State Key Laboratory for Diagnosis and Treatment of Infectious Diseases at Zhejiang University for glycoproteomic analysis. We thank Shuangtian Shengwu Biotech. Co.Ltd for assistance of BIL analysis. We thank pGlyco team at Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, for technical support for pGlyco analysis. We thank Dr. Ma Ping for consultation on protein expression and purification.

Funding

This work was supported by The National Key Research and Development Program (2017YFC1200204, 2017YFA0504803, 2018YFA0507700), Emergency Project of Zhejiang Provincial Department of Science and Technology (2020C03123-1), Fundamental Research Funds for the Central Universities (2018XZZX001-13) and Independent Project Fund of the State Key Laboratory for Diagnosis and Treatment of Infectious Disease.

Authors' contributions:

Study concept and design: ZS, LL; Samples preparation: ZS, KR, FJ, XO; LC-MSMS experiments: ZS, KR; Analysis and interpretation of data: ZS, JJ, KR; Structure Modeling: XZ, JC, ZS; Drafting of the manuscript: ZS, ZJ; Critical revision of the manuscript: LL; All authors approved the final version of the manuscript, including the authorship list.

Competing interests

The authors declare that they have no competing interests.

Figure Legend

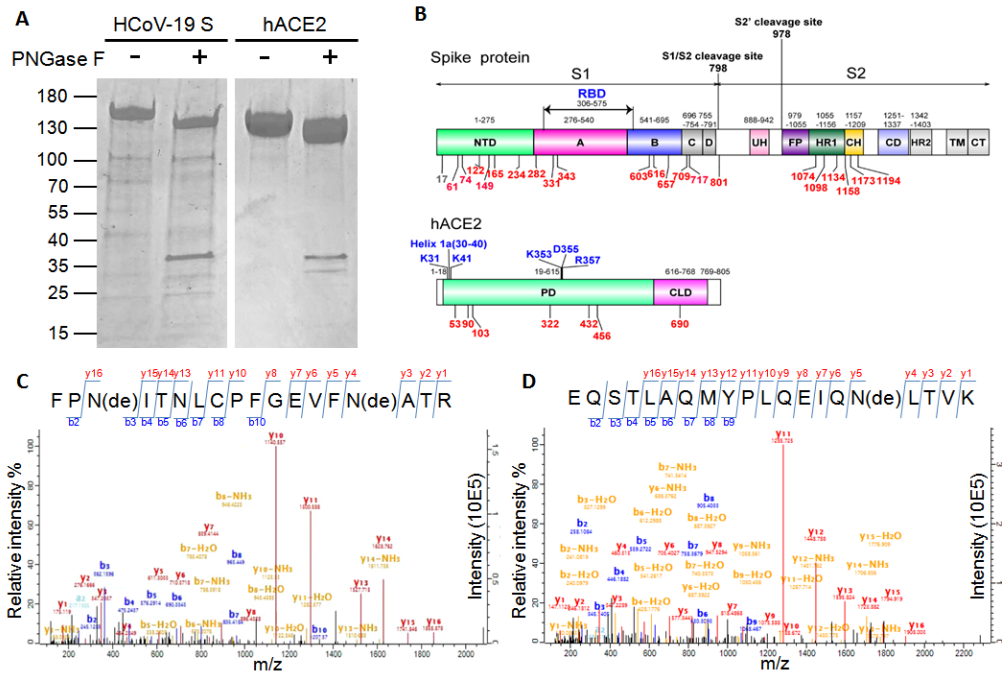


Figure. 1. Potential glycosylation sites in HCoV-19 spike protein and hACE2.

A: 15% SDS-PAGE analysis of intact and deglycosylated form of HCoV-19 spike protein and hACE2. Molecular weight markers are shown on the left. **B:** schematic representation of functional subunits and domains of HCoV-19 spike protein (upper panel) and hACE2 (lower panel). CD, connector domain; CH, central helix; CT, cytoplasmic tail; FP, fusion peptide; TM, transmembrane domain; UH, upstream helix; HR1/2, heptad repeat 1/2. Blue indicated sites possibly responsible for interaction between the S protein and hACE2. Potential glycosylation sites within in each domain were listed in braces. Red indicated identified glycosylated sites in this study. **M**ass spectra of identified deglycopeptide containing N331 and N343 in HCoV-19 spike protein (**C**) and deglycopeptide containing N90 in hACE2 (**D**).

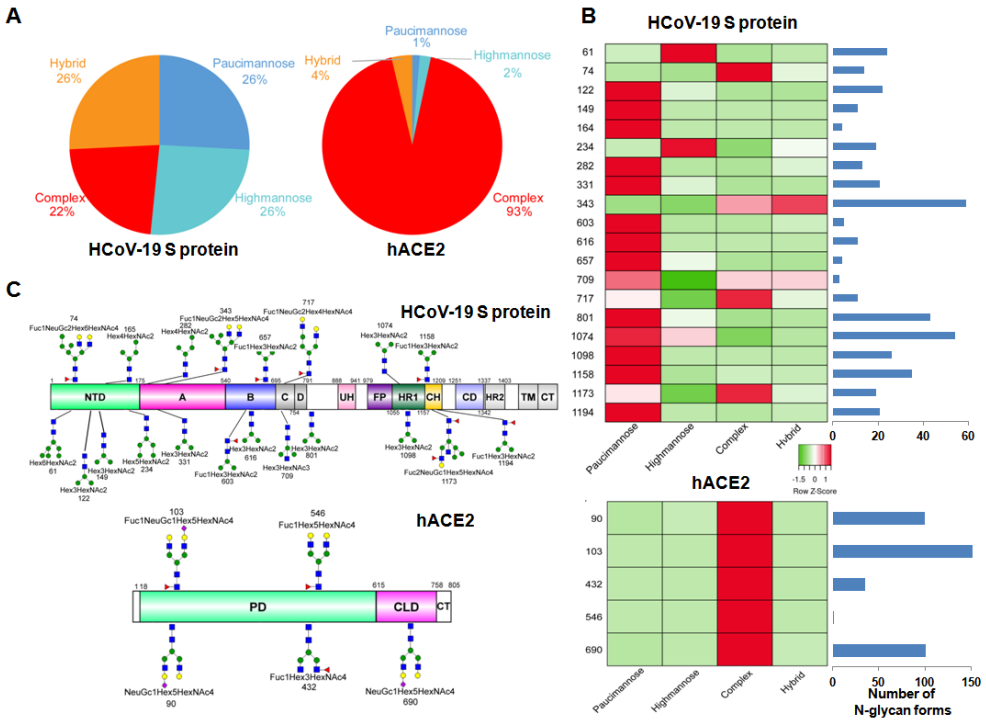


Figure. 2, Summary of N-glycopeptide survey in HCoV-19 Spike protein and hACE2.

A: Overall percentages of 4 major N-glycan categories identified in either protein. **B:** schematic representation of predominant N-glycan forms in each site related to functional domains in either protein. **C:** relative LC-MS intensity of 4 major N-glycan categories identified in each site were presented in heatmap, alongside with total number of identified glycans summarized in bars plots.

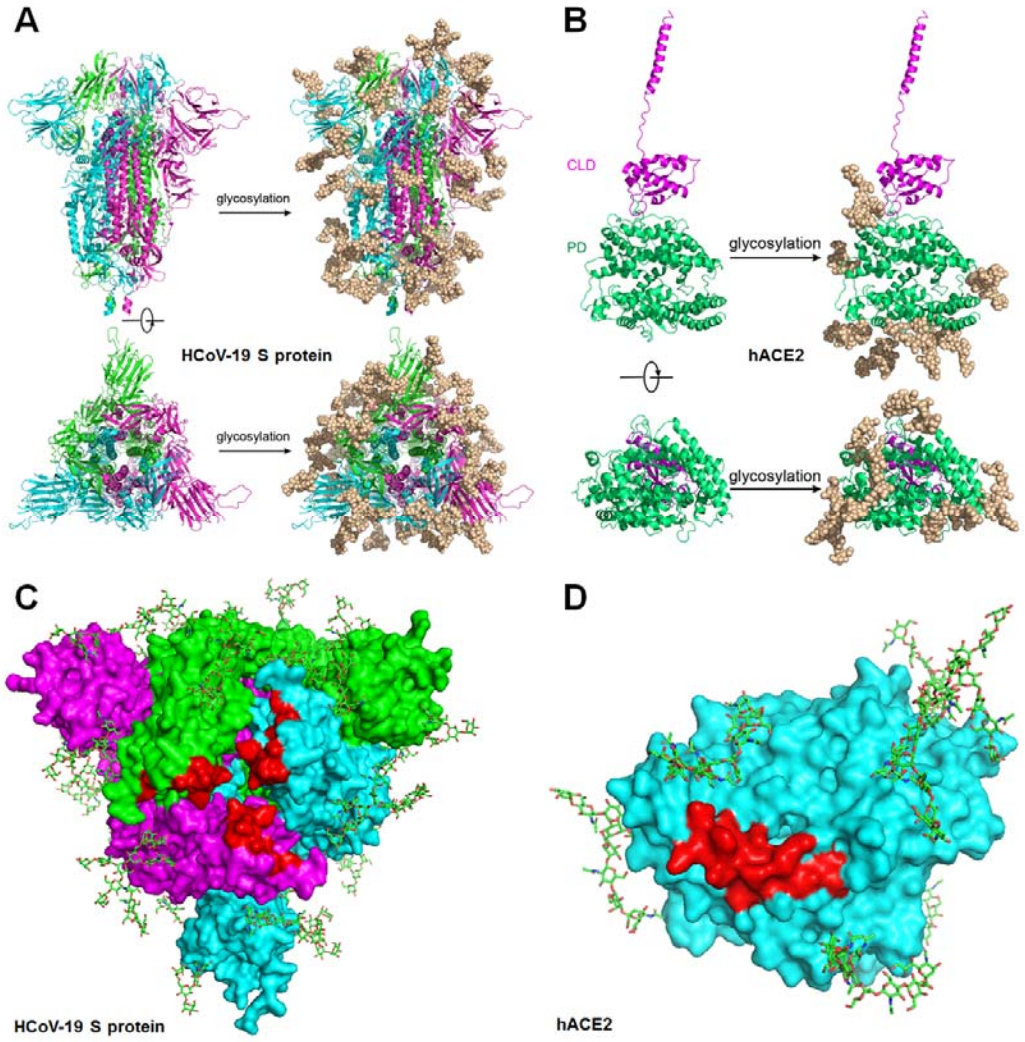


Figure. 3, Refined structure model of HCoV-19 Spike trimer and hACE2 incorporating N-glycans.

3D ribbon diagrams of HCoV-19 spike trimer colored by protomer (A) and hACE2 colored by major domains (B). For both model, side-view (upper panel) and top-view looking towards the viral or cellular membrane (lower panel) are shown, model without (left panel) or with glycans (right panel) were both provided. Diagrams showing binding sites (red) with glycans in vicinity on the surface of HCoV-19 spike trimer (C) and hACE2 (D) in top-view. Glycans were presented in sticks.

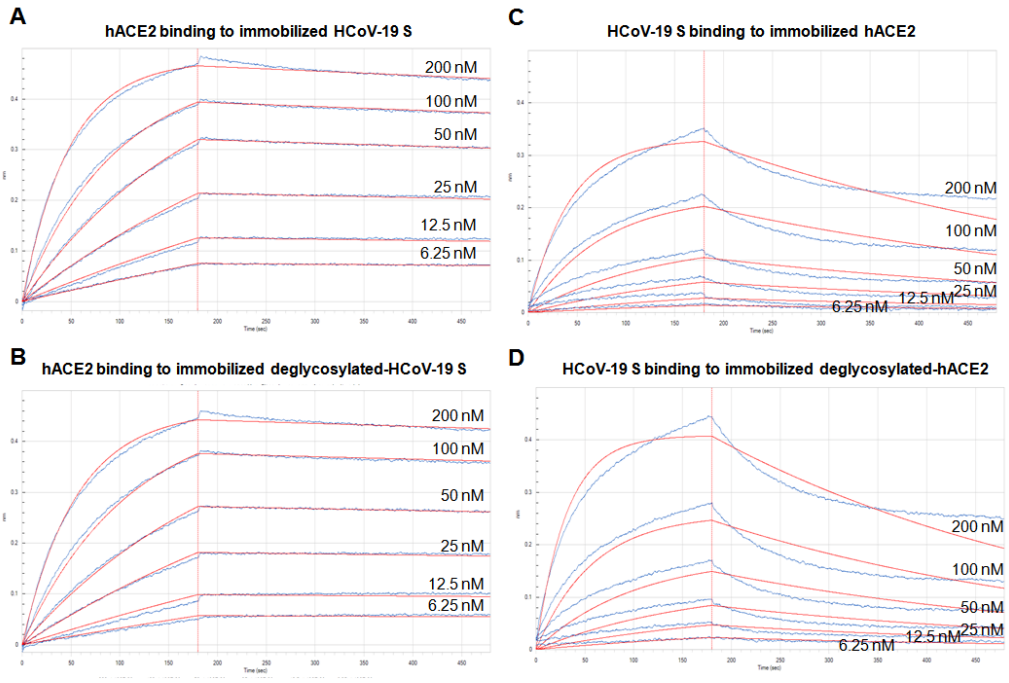


Figure. 4, Impact of glycosylation on binding between HCoV-19 Spike protein and hACE2.

Binding of HCoV-19 S protein and hACE2 was measured by biolayer interferometry. Biosensors with immobilized intact HCoV-19 S protein (A) and its deglycosylated form (B) were exposed to hACE2 at concentration of 12.5, 25, 50, 100, 200 nM. Swap experiments were conducted using biosensors with immobilized intact hACE2 (C) and its deglycosylated form (D) exposed to HCoV-19 S protein at concentration of 12.5, 25, 50, 100, 200 nM. Red lines correspond to a global fit of the data using an equimolar binding model.

Tables

Table 1, Summary of PTM identified in deglycosylated peptides from HCoV-19 S and hACE2 protein.

Sites	PTM	Sequence	Occupancy %
HCoV-19 S			
61;74	N-glycosylation	SSVLHSTQDLFLPFFS**N**#VTWFHAIHVSGT **N**#GTK	100
74	N-glycosylation	HAIHVSGT**N**#GTKRF	100
78	Methylation	**R**#FDNPVLPFNDGVYFASTEK	100
122	N-glycosylation	TQSLLIVN**N**#ATNVVIK IVN**N**#ATNVVIKVCEF	100 100
165	N-glycosylation	VYSSAN**N**#CTFEYVSQPFLMDLEGK	100
224	Methylation	DLPQGFSAL**E**#PLVDLPIGINITR	100
234	N-glycosylation	VDLPIGI**N**#ITRF	100
282	N-glycosylation	YNE**N**#GTITDAVDCALDPLSETK NE**N**#GTITDAVDCALDPLSETKCTL	100 100
331	N-glycosylation	RVQPTESIVRFP**N**#ITNL	100
331;334	N-glycosylation	FP**N**#ITNLCPFGEVF**N**#ATR	100
340	Methylation	FPNITNLCPFG**E**#VFNATR	2.02
603	N-glycosylation	GGVSVITPGT**N**#TSNQVAVL	43.50
616	N-glycosylation	YQDV**N**#CTEVPVAIHADQL	100

654;661	Methylation	AGCLIGA**E**#HVNNSY**E**#CDIPIGAGICASYQ TQTNSPR	100
657	N-glycosylation	AGCLIGAEHV**N**#NSYECDIPIGAGICASYQT QTNSPR IGAEHV**N**#NSYECDIPIGAGICASY	74.80 79.76
709;717	N-glycosylation	S**N**#NSIAIPT**N**#F	100
801	N-glycosylation	TPPIKDFGGF**N**#FSQILPDPSKPSK **N**#FSQILPDPSKPSKRSF	100 100
1074	N-glycosylation	**N**#FTTAPAICHDGK	100
1098	N-glycosylation	VS**N**#GTHWF	100
1134	N-glycosylation	VSGNCDVVIGIV**N**#NTVY	100
1158;1173	N-glycosylation	**N**#HTSPDVDLGDISGI**N**#ASVVNIQKEIDRL NEVAK	99.17
1194	N-glycosylation	NL**N**#ESLIDLQELGK	99.73
hACE2			
53	N-glycosylation	FNHEAEDLFYQSSLASWNYNT**N**#ITEENVQ NMNNAGDK NYNT**N**#ITEENVQNMNNAGDKW	100 100
68	Methylation	NYNTNITEENVQNMNNAGD**K**#W	100
84	Hydroxyproline	EQSTLAQMY**P**#LQEIQNLTVK	20.74
90	N-glycosylation	WSAFLKEQSTLAQMYPLQEIQ**N**#LTVK	100
103	N-glycosyla	LQLQALQQ**N**#GSSVLSEDKSK	100

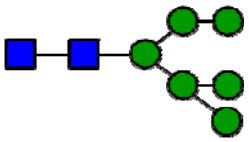
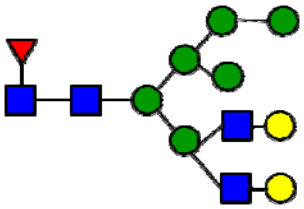
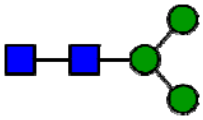
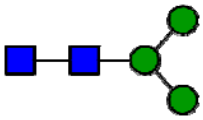
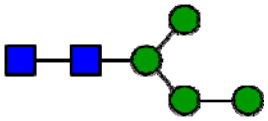
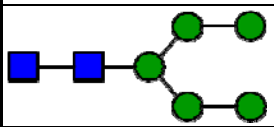
	tion	QLQALQQ**N**#GSSVL	100
253	Hydroxypro line	LMNAY**P**#SYISPIGCLPAHLLGDMWGR	98.35
263	Hydroxypro line	LMNAYPSYISPIGCL**P**#AHLLGDMWGR	88.57
284	Hydroxypro line	FWTNLYSLTV**P**#FGQKPNIDVTDAMVDQA WDAQR	4.06
289	Hydroxypro line	**P**#NIDVTDAMVDQAWDAQR	1.70
321	Hydroxypro line	FFVSVGL**P**#NMTQGFWENSMLTDPGNVQ K	58.24
322	N-glycosyla tion	FFVSVGLP**N**#MTQGFWENSMLTDPGNVQK VSVGLP**N**#MTQGF	100 100
329	Methylation	FFVSVGLPNMTQGFW**E**#NSMLTDPGNVQ K	100
346	Hydroxypro line	AVCH**P**#TAWDLGK	88.88
415	Hydroxypro line	NGANEGFHEAVGEIMSLSAAT**P**#K	0.68
432	N-glycosyla tion	SIGLLSPDFQEDN#ETEINFLLK LSPDFQEDN#ETEINF	100 100
451	Hydroxypro line	QALTIVGTL**P**#FTYMLEK	43.59
469	Hydroxypro line	GEI**P**#KDQWMK	4.09
546	N-glycosyla tion	CDIS**N**#STEAGQK HKCDIS**N**#STEAGQKLF	100 100
565	Hydroxypro	SE**P**#WTLALENVVGAK	0.37

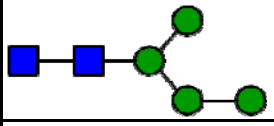
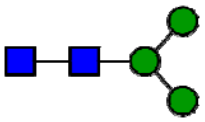
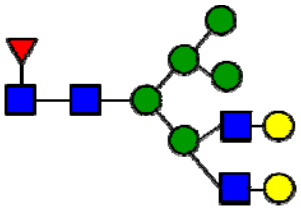
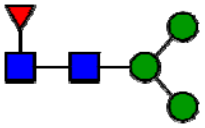
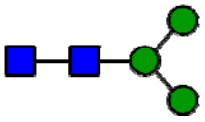
	line		
583	Hydroxyproline	NMNVR**P**[#]LLNYFEPLFTWLK	2.12
612	Hydroxyproline	NSFVGWSTDWS**P**[#]YADQSIK	0.21
690	N-glycosylation	ISFNFFVTAPK**N**[#]VSDIIPR VTAPK**N**[#]VSDIIPRTEVEKAIRM **N**[#]VSDIIPRTEVEK	100 100 100

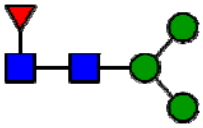
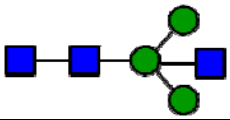
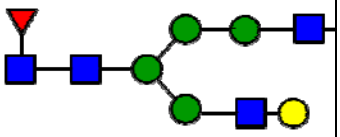
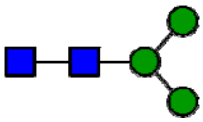
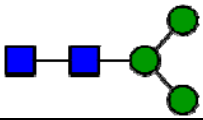
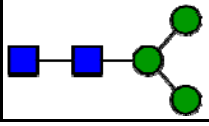
Note: Modified sites were indicated by [#].

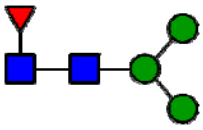
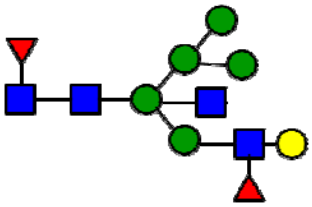
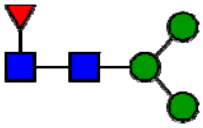
511 **Table 2. Summary of glycopeptide and N-glycan identification in HCoV-19 S**

512 **protein.**

Domain	Site	Sequences	Most abundant composition	Most likely structure
S1	61	FS**N**#VTWF FS**N**#VTW	Hex6HexNAc2	
	74	SSVLHSTQDLFLPFFS**N**#VTWFHAIHVSGT**N**#GTK SSVLHSTQDLFLPFFS**N**#VTWFHAIHVSGT**N**#GTKR HAIHVSGT**N**#GTKRF	Fuc1NeuGc2Hex6HexNAc4	
	122	TQSLLIVN**N**#ATNVVIK IVN**N**#ATNVVIKVCEF LIVN**N**#ATNVVIKVCEF	Hex3HexNAc2	
	149	VCEFQFCNDPFLGVY YHKN**N**#K YHKN**N**#KSWMESEF HKN**N**#KSWMESEF	Hex3HexNAc2	
	165	VYSSAN**N**#CTFEYVSQ PFLMDLEGK SSAN**N**#CTFEY SSAN**N**#CTFEYVSQPF	Hex4HexNAc2	
	234	DLPQGFSALEPLVDLPI GI**N**#ITR	Hex5HexNAc2	

		VDLPIGI**N**#ITRF		
	282	YNE**N**#GTITDAVDCAL DPLSETK	Hex4HexNAc2	
	331	FP**N**#ITNLCPFGEVF**N**# ATR RVQPTESIVRFP**N**#ITN LCPF RVQPTESIVRFP**N**#ITN L	Hex3HexNAc2	
	343	FPNITNLCPFGEVF**N**#A TR CPFGEVF**N**#ATRF GEVF**N**#ATRF	Fuc1NeuGc2He x5HexNAc4	
	603	GGVSVITPGT**N**#TSNQ VAVLY GGVSVITPGT**N**#TSNQ VAVL	Fuc1Hex3HexN Ac2	
	616	YQDV**N**#CTEVPVAIHA DQLTPTW QDV**N**#CTEVPVAIHAD QLTPTW QDV**N**#CTEVPVAIHAD QLTPTWRVY YQDV**N**#CTEVPVAIHA DQL QDV**N**#CTEVPVAIHAD QL	Hex3HexNAc2	

	657	AGCLIGAEHVN#NSYECDIPIGAGICASYQTQTNSPR QTRAGCLIGAEHVN#NSYECDIPIGAGICASY IGAEHVN#NSYECDIPIGAGICASY QTRAGCLIGAEHVN#NSY	Fuc1Hex3HexNAc2	
S2	709	SN#NSIAIPTNF	Hex3HexNAc3	
	717	SNNSIAIPTN#F	Fuc1NeuGc2Hex4HexNAc4	
	801	DFGGFN#FSQILPDPSKPSK DFGGFN#FSQILPDPSKPSKR TPPIKDFGGFN#FSQILPDPSK TPPIKDFGGFN#FSQILPDPSKPSK DFGGFN#FSQILPDPSKN#FSQILPDPSKPSKRSF	Hex3HexNAc2	
S2(S2')	1074	N#FTTAPAICHDGK LHVTYVPAQEKN#F	Hex3HexNAc2	
	1098	EGVFVSN#GTHWFVTQR	Hex3HexNAc2	

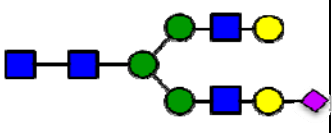
		AHFPREGVFVSN#GTHWFVTQR VSN#GTHWF VSN#GTHW		
	1134*	-	-	-
	1158	N#HTSPDVDLGDISGINASVVNIQK YFKN#HTSPDVDLGDISGINASVVNIQK	Fuc1Hex3HexNAc2	
	1173	YFKNHTSPDVDLGDISGIN#ASVVNIQK NHTSPDVDLGDISGIN#ASVVNIQK	Fuc2NeuGc1Hex5HexNAc4	
	1194	NLN#ESLIDLQELGKYEQYIK NLN#ESLIDLQELGK	Fuc1Hex3HexNAc2	

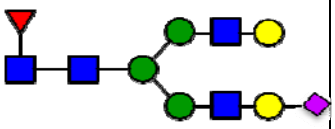
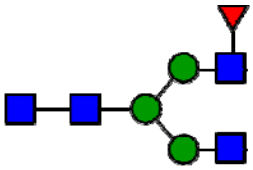
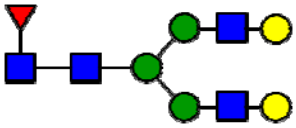
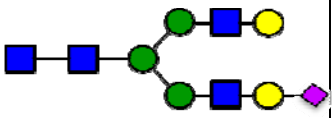
513 Note: glycosylated N sites were indicated by #. Sites identified only on deglycosylated peptides

514 were indicated by *.

515

516 **Table 3, Summary of glycopeptide and N-glycan identification in hACE2 protein.**

Domain	Site	Sequences	Most abundant composition	Most likely structure
PD	53*	-	-	-
	90	EQSTLAQMYPLQEIQN#LTVK	NeuGc1Hex5HexNAc4	

	103	LQLQALQQ**N**#GSSVLSEDK LQLQALQQ**N**#GSSVLSEDKSK LQLQALQQ**N**#GSSVLSEDKSKR	Fuc1NeuGc1Hex5HexNAc4	
	322*	-	-	-
	423	SIGLLSPDFQED**N**#ETEINFLLK	Fuc1Hex3HexNAc4	
	546	CDIS**N**#STEAGQK	Fuc1Hex5HexNAc4	
CLD	690	**N**#VSDIIPR ISFNFFVTAPK**N**#VSDIIPR	NeuGc1Hex5HexNAc4	

Note: glycosylated N sites were indicated by #. Sites identified only on deglycosylated peptides were indicated by *.

References:

Biswas, A., Bhattacharjee, U., Chakrabarti, A.K., Tewari, D.N., Banu, H., and Dutta, S. (2020). Emergence of Novel Coronavirus and COVID-19: whether to stay or die out? Critical reviews in microbiology, 1-12.

Chang, D., and Zaia, J. (2019). Why Glycosylation Matters in Building a Better Flu Vaccine. Molecular & cellular proteomics : MCP *18*, 2348-2358.

Chen, W.H., Du, L., Chag, S.M., Ma, C., Tricoche, N., Tao, X., Seid, C.A., Hudspeth, E.M., Lustigman, S., Tseng, C.T., *et al.* (2014). Yeast-expressed recombinant protein of the receptor-binding domain in

529 SARS-CoV spike protein with deglycosylated forms as a SARS vaccine candidate. Human vaccines &
530 immunotherapeutics *10*, 648-658.
531 Cheng, Y., Cheng, G., Chui, C.H., Lau, F.Y., Chan, P.K., Ng, M.H., Sung, J.J., and Wong, R.S. (2005). ABO
532 blood group and susceptibility to severe acute respiratory syndrome. Jama *293*, 1450-1451.
533 de Groot, R.J. (2006). Structure, function and evolution of the hemagglutinin-esterase proteins of
534 corona- and toroviruses. Glycoconjugate journal *23*, 59-72.
535 Dong, E., Du, H., and Gardner, L. (2020). An interactive web-based dashboard to track COVID-19 in real
536 time. The Lancet Infectious diseases.
537 Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta
538 crystallographica Section D, Biological crystallography *66*, 486-501.
539 Feldmann, H., Nichol, S.T., Klenk, H.D., Peters, C.J., and Sanchez, A. (1994). Characterization of
540 filoviruses based on differences in structure and antigenicity of the virion glycoprotein. Virology *199*,
541 469-473.
542 Fukushi, M., Yoshinaka, Y., Matsuoka, Y., Hatakeyama, S., Ishizaka, Y., Kirikae, T., Sasazuki, T., and
543 Miyoshi-Akiyama, T. (2012). Monitoring of S protein maturation in the endoplasmic reticulum by
544 calnexin is important for the infectivity of severe acute respiratory syndrome coronavirus. Journal of
545 virology *86*, 11745-11753.
546 Guan, W.J., Ni, Z.Y., Hu, Y., Liang, W.H., Ou, C.Q., He, J.X., Liu, L., Shan, H., Lei, C.L., Hui, D.S.C., *et al.*
547 (2020). Clinical Characteristics of Coronavirus Disease 2019 in China. The New England journal of
548 medicine.
549 Guillon, P., Clement, M., Sebille, V., Rivain, J.G., Chou, C.F., Ruvoen-Clouet, N., and Le Pendu, J. (2008).
550 Inhibition of the interaction between the SARS-CoV spike protein and its cellular receptor by
551 anti-histo-blood group antibodies. Glycobiology *18*, 1085-1093.
552 Han, D.P., Lohani, M., and Cho, M.W. (2007). Specific asparagine-linked glycosylation sites are critical
553 for DC-SIGN- and L-SIGN-mediated severe acute respiratory syndrome coronavirus entry. Journal of
554 virology *81*, 12029-12039.
555 Hoffmann, M., Kleine-Weber, H., Schroeder, S., Kruger, N., Herrler, T., Erichsen, S., Schiergens, T.S.,
556 Herrler, G., Wu, N.H., Nitsche, A., *et al.* (2020). SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2
557 and Is Blocked by a Clinically Proven Protease Inhibitor. Cell.
558 Jeffers, S.A., Tusell, S.M., Gillim-Ross, L., Hemmila, E.M., Achenbach, J.E., Babcock, G.J., Thomas, W.D.,

559 Jr., Thackray, L.B., Young, M.D., Mason, R.J., *et al.* (2004). CD209L (L-SIGN) is a receptor for severe
560 acute respiratory syndrome coronavirus. Proceedings of the National Academy of Sciences of the
561 United States of America *101*, 15748-15753.
562 Krokhin, O., Li, Y., Andonov, A., Feldmann, H., Flick, R., Jones, S., Stroeher, U., Bastien, N., Dasuri, K.V.,
563 Cheng, K., *et al.* (2003). Mass spectrometric characterization of proteins from the SARS virus: a
564 preliminary report. Molecular & cellular proteomics : MCP *2*, 346-356.
565 Kumar, S., Maurya, V.K., Prasad, A.K., Bhatt, M.L.B., and Saxena, S.K. (2020). Structural, glycosylation
566 and antigenic variation between 2019 novel coronavirus (2019-nCoV) and SARS coronavirus
567 (SARS-CoV). Virusdisease *31*, 13-21.
568 Li, W., Hulswit, R.J.G., Widjaja, I., Raj, V.S., McBride, R., Peng, W., Widagdo, W., Tortorici, M.A., van
569 Dieren, B., Lang, Y., *et al.* (2017). Identification of sialic acid-binding function for the Middle East
570 respiratory syndrome coronavirus spike glycoprotein. Proceedings of the National Academy of
571 Sciences of the United States of America *114*, E8508-E8517.
572 Liu, J., Cao, R., Xu, M., Wang, X., Zhang, H., Hu, H., Li, Y., Hu, Z., Zhong, W., and Wang, M. (2020).
573 Hydroxychloroquine, a less toxic derivative of chloroquine, is effective in inhibiting SARS-CoV-2
574 infection in vitro. Cell discovery *6*, 16.
575 Liu, M.Q., Zeng, W.F., Fang, P., Cao, W.Q., Liu, C., Yan, G.Q., Zhang, Y., Peng, C., Wu, J.Q., Zhang, X.J., *et*
576 *al.* (2017). pGlyco 2.0 enables precision N-glycoproteomics with comprehensive quality control and
577 one-step mass spectrometry for intact glycopeptide identification. Nature communications *8*, 438.
578 Parsons, L.M., Bouwman, K.M., Azurmendi, H., de Vries, R.P., Cipollo, J.F., and Verheije, M.H. (2019).
579 Glycosylation of the viral attachment protein of avian coronavirus is essential for host cell and
580 receptor binding. The Journal of biological chemistry *294*, 7797-7809.
581 Raman, R., Tharakaraman, K., Sasisekharan, V., and Sasisekharan, R. (2016). Glycan-protein
582 interactions in viral pathogenesis. Current opinion in structural biology *40*, 153-162.
583 Ritchie, G., Harvey, D.J., Feldmann, F., Stroeher, U., Feldmann, H., Royle, L., Dwek, R.A., and Rudd, P.M.
584 (2010). Identification of N-linked carbohydrates from severe acute respiratory syndrome (SARS) spike
585 glycoprotein. Virology *399*, 257-269.
586 Robinson, B.S., Arthur, C.M., Evavold, B., Roback, E., Kamili, N.A., Stowell, C.S., Vallecillo-Zuniga, M.L.,
587 Van Ry, P.M., Dias-Baruffi, M., Cummings, R.D., *et al.* (2019). The Sweet-Side of Leukocytes: Galectins
588 as Master Regulators of Neutrophil Function. Frontiers in immunology *10*, 1762.

589 Shih, Y.P., Chen, C.Y., Liu, S.J., Chen, K.H., Lee, Y.M., Chao, Y.C., and Chen, Y.M. (2006). Identifying
590 epitopes responsible for neutralizing antibody and DC-SIGN binding on the spike glycoprotein of the
591 severe acute respiratory syndrome coronavirus. Journal of virology *80*, 10315-10324.
592 Song, H.C., Seo, M.Y., Stadler, K., Yoo, B.J., Choo, Q.L., Coates, S.R., Uematsu, Y., Harada, T., Greer, C.E.,
593 Polo, J.M., *et al.* (2004). Synthesis and characterization of a native, oligomeric form of recombinant
594 severe acute respiratory syndrome coronavirus spike glycoprotein. Journal of virology *78*,
595 10328-10335.
596 Sun, Z., Liu, X., Jiang, J., Huang, H., Wang, J., Wu, D., and Li, L. (2016). Toward Biomarker Development
597 in Large Clinical Cohorts: An Integrated High-Throughput 96-Well-Plate-Based Sample Preparation
598 Workflow for Versatile Downstream Proteomic Analyses. Analytical chemistry *88*, 8518-8525.
599 Vigerust, D.J., and Shepherd, V.L. (2007). Virus glycosylation: role in virulence and immune interactions.
600 Trends in microbiology *15*, 211-218.
601 Vincent, M.J., Bergeron, E., Benjannet, S., Erickson, B.R., Rollin, P.E., Ksiazek, T.G., Seidah, N.G., and
602 Nichol, S.T. (2005). Chloroquine is a potent inhibitor of SARS coronavirus infection and spread. Virology
603 journal *2*, 69.
604 Walls, A.C., Park, Y.J., Tortorici, M.A., Wall, A., McGuire, A.T., and Veesler, D. (2020a). Structure,
605 Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. Cell.
606 Walls, A.C., Park, Y.J., Tortorici, M.A., Wall, A., McGuire, A.T., and Veesler, D. (2020b). Structure,
607 Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. Cell *181*, 281-292 e286.
608 Wan, Y., Shang, J., Graham, R., Baric, R.S., and Li, F. (2020). Receptor Recognition by the Novel
609 Coronavirus from Wuhan: an Analysis Based on Decade-Long Structural Studies of SARS Coronavirus.
610 Journal of virology *94*.
611 Wang, D., Hu, B., Hu, C., Zhu, F., Liu, X., Zhang, J., Wang, B., Xiang, H., Cheng, Z., Xiong, Y., *et al.* (2020).
612 Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus-Infected Pneumonia
613 in Wuhan, China. Jama.
614 Wang, W.H., Lin, C.Y., Chang, M.R., Urbina, A.N., Assavalapsakul, W., Thitithanyanont, A., Chen, Y.H., Liu,
615 F.T., and Wang, S.F. (2019). The role of galectins in virus infection - A systemic literature review. Journal
616 of microbiology, immunology, and infection = Wei mian yu gan ran za zhi.
617 Watanabe, Y., Bowden, T.A., Wilson, I.A., and Crispin, M. (2019). Exploitation of glycosylation in
618 enveloped virus pathobiology. Biochimica et biophysica acta General subjects *1863*, 1480-1497.

Wrapp, D., Wang, N., Corbett, K.S., Goldsmith, J.A., Hsieh, C.L., Abiona, O., Graham, B.S., and McLellan, J.S. (2020). Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science *367*, 1260-1263.

Wu, A., Peng, Y., Huang, B., Ding, X., Wang, X., Niu, P., Meng, J., Zhu, Z., Zhang, Z., Wang, J., *et al.* (2020a). Genome Composition and Divergence of the Novel Coronavirus (2019-nCoV) Originating in China. Cell host & microbe *27*, 325-328.

Wu, F., Zhao, S., Yu, B., Chen, Y.M., Wang, W., Song, Z.G., Hu, Y., Tao, Z.W., Tian, J.H., Pei, Y.Y., *et al.* (2020b). A new coronavirus associated with human respiratory disease in China. Nature *579*, 265-269.

Xu, Y. (2020). Unveiling the Origin and Transmission of 2019-nCoV. Trends in microbiology *28*, 239-240.

Yang, T.J., Chang, Y.C., Ko, T.P., Draczkowski, P., Chien, Y.C., Chang, Y.C., Wu, K.P., Khoo, K.H., Chang, H.W., and Hsu, S.D. (2020). Cryo-EM analysis of a feline coronavirus spike protein reveals a unique structure and camouflaging glycans. Proceedings of the National Academy of Sciences of the United States of America *117*, 1438-1446.

Ying, W., Hao, Y., Zhang, Y., Peng, W., Qin, E., Cai, Y., Wei, K., Wang, J., Chang, G., Sun, W., *et al.* (2004). Proteomic analysis on structural proteins of Severe Acute Respiratory Syndrome coronavirus. Proteomics *4*, 492-504.

York, I.A., Stevens, J., and Alymova, I.V. (2019). Influenza virus N-linked glycosylation and innate immunity. Bioscience reports *39*.

Zhang, T., Wu, Q., and Zhang, Z. (2020). Probable Pangolin Origin of SARS-CoV-2 Associated with the COVID-19 Outbreak. Current biology : CB.

Zhang, Y.Z., and Holmes, E.C. (2020). A Genomic Perspective on the Origin and Emergence of SARS-CoV-2. Cell.

Zhao, J., Yang, Y., Huang, H., Li, D., Gu, D., Lu, X., Zhang, Z., Liu, L., Liu, T., Liu, Y., *et al.* (2020). Relationship between the ABO Blood Group and the COVID-19 Susceptibility. medRxiv.

Zhao, X., Guo, F., Comunale, M.A., Mehta, A., Sehgal, M., Jain, P., Cuconati, A., Lin, H., Block, T.M., Chang, J., *et al.* (2015). Inhibition of endoplasmic reticulum-resident glucosidases impairs severe acute respiratory syndrome coronavirus and human coronavirus NL63 spike protein-mediated entry by altering the glycan processing of angiotensin I-converting enzyme 2. Antimicrobial agents and chemotherapy *59*, 206-216.

Zheng, J., Yamada, Y., Fung, T.S., Huang, M., Chia, R., and Liu, D.X. (2018). Identification of N-linked

649 glycosylation sites in the spike protein and their functional impact on the replication and infectivity of
650 coronavirus infectious bronchitis virus in cell culture. Virology *513*, 65-74.
651 Zheng, L., Li, H., Fu, L., Liu, S., Yan, Q., and Leng, S.X. (2020). Blocking cellular N-glycosylation
652 suppresses human cytomegalovirus entry in human fibroblasts. Microbial pathogenesis *138*, 103776.
653 Zhou, F., Yu, T., Du, R., Fan, G., Liu, Y., Liu, Z., Xiang, J., Wang, Y., Song, B., Gu, X., *et al.* (2020a). Clinical
654 course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a
655 retrospective cohort study. Lancet.
656 Zhou, P., Yang, X.L., Wang, X.G., Hu, B., Zhang, L., Zhang, W., Si, H.R., Zhu, Y., Li, B., Huang, C.L., *et al.*
657 (2020b). A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature *579*,
658 270-273.
659 Zhou, Y., Lu, K., Pfefferle, S., Bertram, S., Glowacka, I., Drosten, C., Pohlmann, S., and Simmons, G.
660 (2010). A single asparagine-linked glycosylation site of the severe acute respiratory syndrome
661 coronavirus spike glycoprotein facilitates inhibition by mannose-binding lectin through multiple
662 mechanisms. Journal of virology *84*, 8753-8764.