

N-glycoproteomic Profiling Revealing New Coronavirus Therapeutic Targets That Maybe Involved in Cepharranthine's Intervention

Wenlin An (✉ wenlin.an@mail.buct.edu.cn)

College of Life Science and Technology, Beijing University of Chemical Technology

Fengjuan Tian

College of Life Science and Technology, Beijing University of Chemical Technology

Jing Li

College of Life Science and Technology, Beijing University of Chemical Technology

Yanglan Wang

College of Life Science and Technology, Beijing University of Chemical Technology

Shihui Wang

College of Life Science and Technology, Beijing University of Chemical Technology

Huahao Fan

College of Life Science and Technology, Beijing University of Chemical Technology

Yigang Tong

College of Life Science and Technology, Beijing University of Chemical Technology

Article

Keywords: N-glycosylation, COVID-19, SARS-CoV-2, coronavirus, vaccine development, viral pathogenesis

Posted Date: September 24th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-75332/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

N-glycosylation is an important post-translational modification involved in protein folding, signal transduction, extracellular matrix organization and immune response. Evidence showed that glycosylated SARS-CoV-2 Spike protein may be a potential target in viral pathogenesis and drug/vaccine design. To investigate the mechanism of coronavirus infestation and drug targets from glycosylation perspective, we constructed a SARS-CoV-2 cellular model using GX_P2V-infected VeroE6 cells to study the effects of GX_P2V on glycoproteins in presence or absence of Cepharranthine (CEP) through N-glycoproteomics profiling. The results showed that coronavirus GX_P2V could cause aberrant protein glycosylation, whereas CEP can partially maintain GX_P2V-induced aberrant glycoproteins at homeostasis. Further study revealed that proteins LAMB1 and FN1 were pivotal in counteracting coronavirus-induced aberrant protein glycosylation by CEP. Furthermore, CEP can dramatically regulate the glycosylation of viral proteins S, M and N. Our results suggest that despite the strong anti-coronavirus effects of CEP, drug combinations need be considered to achieve optimal therapeutic strategies.

Introduction

The current global pandemic of COVID-19 caused by the highly contagious coronavirus SARS-CoV-2 has had a huge impact on international community. Currently, no anti-SARS-CoV-2 drugs or vaccines have been officially and clinically approved due to the absence of adequate evidence for exact mechanism of SARS-COV-2 infestation. Therefore, it is of great urgent need to deeply understand its mechanism to screen potential drug targets¹.

Glycosylation is the most abundant and complex post-translational modification which can have profound structural and functional effects on the conjugate². The oligosaccharide fraction is considered to be related with multiple biological processes and affects the physical properties of proteins. It has been known that protein glycosylation is definitely stable among the normal people, although it is different from person to person. Only when a person's homeostasis changes due to lifestyle or pathological conditions, will glycosylation changes significantly^{3,4}. The glycosylation has been labeled an important quality attribute of biopharmaceuticals⁵. The S1 N-terminal domain of the spike protein (S1-NTD) may "hide" sites for recognition and binding of sialic acid-containing glycans, which may provide possibilities for ACE2 alternative receptors⁶. Therefore, it is reasonable to take the glycosylation in consideration in the treatment of coronavirus infestation.

GX_P2V is a SARS-CoV-2 related coronavirus isolated from pangolin. Its spike protein shares 92.2% amino acid identity with the spike protein of SARS-CoV-2⁷. Besides, GX_P2V also has the shared receptor with SARS-CoV-2 but has no pathogenicity to human. Based on the above characteristics, it has been reported that GX_P2X can be used as an accessible in vitro model for developing therapies or for mechanism investigation against SARS-CoV-2⁷. Here we adopted this well-established coronavirus SARS-CoV-2 cellular model by using GX_P2V infestation with Vero E6 to investigate the mechanisms of

coronavirus SARS-CoV-2 infestation with respect to the dysregulation of glycoprotein and screening for potential targets for therapeutic intervention.

We have reported that cepharanthine (CEP), a clinically approved drug to increase immune function after chemotherapy, has anti-coronavirus properties that prevent the infected host cells from further damage⁷. But the mechanism of CEP on glycosylation targets remains unclear. In this paper, we carried out N-glycosylation modification proteomics analysis to study the mechanisms of SARS-CoV-2 infection and the anti-coronavirus effects of CEP. Using the advanced technique 4D label-free quantification technology, we constructed the N-glycoproteomic profiling, which revealed new therapeutic targets against coronavirus infestation and CEP's intervention. Our results provide important insights into the mechanism of coronavirus infection and the identification of drug targets for anti-coronavirus of multi-target drug combination usage.

Results And Discussion

N-glycoproteomic profiling highlights the differentially expressed glycoproteins in SARS-CoV-2 infection cellular model and CEP intervention.

N-Glycosylation is one of the most important post-translational modifications of proteins⁸. Although there are only a few hundreds of glycoproteins in total, their highly complex glycosylation increases the number of theoretical protein morphology by several orders of magnitude⁹. Many cell surface and extracellular matrix-related proteins become glycosylated through secretory pathways, thereby regulating their adhesion interactions, physical and chemical properties, and diversifying their functions¹⁰.

To systematically study the differentially expressed glycoproteins during SARS-CoV-2 infestation, we adopted our well-established SARS-CoV-2 cellular model⁷ with GX_P2V-treated Vero E6 cells, and performed the N-glycoproteomic profiling via the 4D label-free LC-MS/MS analysis in presence or absence of GX_P2V and/or CEP. The workflow for the experiment was outlined in supplementary Figure S1. First, the cell samples were harvested and enzymatically digested and then analyzed by 4D label-free LC-MS/MS and bioinformatics. A total of 737,202 secondary spectrograms were obtained by mass spectrometry (supplementary Figure S1B). A total of 1770.0 N-glycosylation modification sites on 828.0 proteins were identified with database searching, of which 1298.0 sites on 624.0 proteins have quantitative information. The detected N-glycosylation tends to obey N-X-T, N-C-S, N-G-S, N-X-S, N-X-C, N-N rules (supplementary Figure S1C).

We then performed the analyses of global proteomics and respective of glycosylation occupancy (Figure 1A), and took the fold changes (F_c = expression levels of glycosylated peptide in Group2 (or 3 or 4)/Group 1 ratios) at $F_c \geq 2$ folds (up-regulated) or $F_c \leq 0.5$ folds (down-regulated) as differentially expressed protein, where group 1 was the mock control (Vero), group 2 was treated with CEP only (Vero_C), group 3 was treated by coronavirus GX_P2V only (Vero_P) and group 4 was treated by CEP and GX_P2V (Vero_C_P). In comparison with the mock control group, GX_P2V treatment can up-regulate 26

glycosylation sites in 23 proteins and down-regulate 58 glycosylation sites in 47 proteins, whereas in presence of CEP, GX_P2V treatment can induce 26 glycosylation sites up-regulated in 22 proteins and 110 glycosylation sites down-regulated in 75 proteins in host Vero E6 cells (Figure 1B).

To investigate the protein-protein interaction (PPI) networks and the interaction modules of these differentially expressed proteins in Vero E6 cells in response to GX_P2V treatment, we conducted STRING and Cytoscape analysis (Figure 1C). We found that the key interaction module obtained from PPI network with MCODE plugin involved in 9 aberrant glycoproteins including 4 up-regulated glycoproteins namely Integrin Subunit Alpha 1 (ITGA1), Integrin Subunit Beta 3 (ITGB3), Laminin Subunit Gamma 1 (LAMC1) and Heparan Sulfate Proteoglycan 2 (HSPG2), and 4 down-regulated glycoproteins namely Integrin Subunit Beta 1 (ITGB1), Laminin Subunit Beta 1 (LAMB1), Integrin Subunit Alpha 5 (ITGA5) and Fibronectin 1 (FN1). The protein (ITGA3) contains both up- and down-regulated glycosylation sites (Figure 1D). LAMC1 was reported to be involved in cell proliferation, angiogenesis, growth, migration and invasion^{11, 12}. Remarkably, majority of the 9 aberrant glycoproteins in the key interaction module were enriched in ECM-receptor interaction pathways (Figure 1D and 1E). Specifically, KEGG pathway enrichment showing GX_P2V infection affected glycoproteins in ECM-receptor interaction (supplementary Figure S2A). ECM components including collagens, laminins, and fibronectins are the major ligands that bind and activate integrin receptors. It helps in mediating cell-cell interactions by binding additional cell receptors or other soluble molecules. ECM plays an important role in cell adhesion, and viral infection may cause significant changes in cell adhesion. Functional enrichment analysis of the key modules revealed that these modules fall into the functional interaction categories of ECM-receptor interaction, laminin interactions, integrin cell surface interactions, endoderm development and receptor-mediated endocytosis. The heavily affected ECM and cell surface interactions might give a reasonable explanation on the gross glass-like changes in the clinical lung image.

Coronavirus GX_P2V-induced alteration in protein glycosylation are prevalent in infected cells

N-linked glycosylation was involved in multiple biological processes, such as receptor interactions, immune responses, protein secretion and transport, cell adhesion, signal transduction, etc. As described above, we detected 84 sites of aberrant glycosylation sites present in 69 glycoproteins from coronavirus GX_P2V-treated Vero E6 cells. To study these sites holistically, we performed functional enrichment clustering analysis of these differential glycoproteins using Metascape (Figure 2A), and the results were graphically presented using the Sangerbox mapping toolbox (Figures 2B and 2C).

Upon the treatment of GX_P2V, the involved functions alteration related to differentially expressed proteins were present in Figure 2A and 2B. The most involved top 20 functions were listed in Figure 2A, and corresponding glycosylation counts of each function presented in Figure 2B. The circle diagram of protein function enrichment analysis demonstrated that ITGB1 was involved in the most relevant function and the extracellular matrix organization was enriched for the most genes. Besides, in response to coronavirus GX_P2V infection, the top 5 aberrant glycoproteins most involved relevant functions include ITGB1, TGFB2, ITGB3, ITGB5 and FN1, and the top 5 most aberrant functions involved most genes were

those related to extracellular matrix organization, response to wounding, regulation of cell adhesion, signaling pathway regulated exocytosis and transmembrane receptor protein tyrosine kinase. Other important functions such as protein maturation, viral entry into host cell and a platelet degranulation are also affected.

To have an overview about the effects of coronavirus GX_P2V treatment on host cells, we summarized the subcellular localization of differentially expressed glycosylated proteins and PPI networks by UniProt database¹³ and STRING software (Figure 3).

Aberrant alteration of protein glycosylation can be seen in cytoplasm, endoplasmic reticulum, Golgi apparatus, lysosome, nucleus, cell membrane and extracellular matrix. The PPI networks of differentially expressed glycoproteins in each subcellular distribution were also highlighted (Figure 3). These results suggest that coronavirus-induced alterations in glycosylation of host cell proteins are widespread in infected cells and have multiple subcellular localizations. The most involved glycoproteins locate in cell membrane, extracellular matrix, lysosomes, and endoplasmic reticulum system and cytosol.

Effects of CEP on GX_P2V-induced aberrant glycosylation of proteins in Vero E6 cells.

We have reported that CEP could effectively inhibit coronavirus infestation⁷. To identify the potential targets of CEP in coronavirus-infected cells in response to GX_P2V treatment, we compared the aberrant glycosylation sites and glycoproteins with and without coronavirus treatment in presence or absence of CEP and perform Reactome Pathway Analysis (Figure 4).

The Venn diagram showed the effect of CEP on the aberrant glycosylation sites of proteins induced by GX_P2V (Figure 4A). Coronavirus GX_P2V treatment can cause abnormal glycosylation at 84 sites (on 69 proteins) (Vero_P/Vero ratio, $F_c \geq 2$ folds), whereas the intervention of CEP can keep 14 glycosylation sites (on 12 glycoproteins) out of 84 GX_P2V-induced aberrant glycosylation sites at normal state (Vero_C_P/Vero ratio, $0.667 \leq F_c \leq 1.5$ folds) (Supplementary Table S1). For the subcellular location of these CEP-affected proteins, aspartate beta-hydroxylase (ASPH) locates in endoplasmic reticulum; glucosylceramidase Beta (GBA) and glucosamine (N-acetyl)-6-sulfatase (GNS) are lysosomal proteins; zinc finger C4H2-type containing (KIAA1166, also named MYORG) is in nucleus ; nicastrin (NCSTN), lysosomal associated membrane protein 1 (LAMP1), protein tyrosine kinase 7 (PTK7) and mannose receptor C type 2 (MRC2) and ATPase Na⁺/K⁺ transporting subunit beta 1 (ATP1B1) are related to cell membrane; laminin subunit beta 1 (LAMB1), fibronectin 1 (FN1) and transforming growth factor beta 2 (TGFB2) are the key proteins related to extracellular matrix(Figure 3 and Figure 4, Supplementary Table S1).

We then carried out Reactome Pathway Analysis on the CEP-affected 12 glycoproteins to understand intervention mechanism of CEP (Figure 4B). The results showed that CEP can exert its antiviral and cell-protective effects by targeting multiple pathways such as ECM proteoglycans, extracellular matrix organization, platelet degranulation, response to elevated platelet, neutrophil degranulation, and etc. Among these 12 proteins, LAMB1 and FN1 are also the pivotal proteins in the key interaction module

which consists of 9 aberrant glycoproteins in response to coronavirus GX_P2V infestation (Figure 1D), indicating that proteins LAMB1 and FN1 might be the key targets of CEP against GX_P2V infestation (Figure 3).

In addition, several proteins can be over-regulated by CEP against GX_P2V (Supplementary Table S1). There are two aberrant glycosylation sites on MRC2 induced by GX_P2V infection. In presence of CEP, the site N69 on MRC2 protein can remain normally glycosylation state, while the other site N1134 is shifted from virus-induced up-regulation to down-regulation. MRC2 may contribute to cellular uptake, remodeling and degradation of extracellular collagen matrices. Like the glycosylation regulation at N1134 site on MRC2 regulated by CEP against GX_P2V, the glycosylation of N83 site on plasminogen activator urokinase receptor (PLAUR) can be over-regulated by CEP, but it is shifted from down-regulation to up-regulation, by which CEP might regulate plasminogen activation system controlling the extracellular level of PLAUR/PLAU, and thus may regulate protease activity at the cell surface^{14,15}. Over-regulated glycosylation by CEP can also be seen at the site N277 of granulin precursor (GRN), which is a secreted glycoprotein that acts as a key regulator of lysosomal function and as a growth factor involved in inflammation, wound healing and cell proliferation¹⁶. It involves in regulating protein trafficking to lysosomes and, also the activity of lysosomal enzymes^{16,17}. Alteration of glycosylation GRN may contribute to lysosome-related functions. Therefore, CEP might also perform its anti-viral effect via regulation of glycosylation of MRC2, PLAUR, and GRN.

Effect of CEP on the N-glycosylation of GX_P2V viral proteins.

Coronavirus viral proteins are normally extensively glycosylated, especially coronavirus spike proteins where it encodes around 66-87 N-linked glycosylation sites per trimeric spike. It has been known that the extensive glycosylation of viral protein plays a role in protein folding and shielding immunogenic epitopes, resulting in immune evasion^{18,19,20}. Remarkably, SARS-CoV and SARS-CoV-2 recognize the human angiotensin converting enzyme-2 (ACE-2) receptor via their glycosylated spike proteins²¹.

From the dataset for N-glycoproteomic profiling, we observed that the identified viral proteins include M, N and Spike were highly glycosylated (Figure 5A), which might offer the viral proteins a glycan shield^{18,19}. Interestingly, we found that the glycosylation of viral proteins M, N and S was affected by CEP treatment. In particular, CEP can up-regulate the glycosylation of the N5 site of M protein, the N-terminal sites of the viral S protein (including the N30, N122, N331 and N343 sites) and the sites N1070 and N1094 sites in the S2 region (N1070 and N1094, the corresponding positions on SARS-CoV-2 spike protein are N1074 and N1098, respectively). However, CEP can dramatically decrease N25-site glycosylation of viral N protein, but no obvious effect can be seen in the glycosylation of the C-terminal sites N1154, N1169, and N1190 of S proteins under CEP treatment.

To observe and analyze the location of these affected glycosylation sites on S proteins and their possible roles, we compared the sequences of spike protein of SARS-CoV-2 and GX_P2V viruses. Due to the structure of GX_P2V viral S protein is not available by far, and considering the high sequence similarity

between SARS-CoV-2 and GX_P2V, we adopted the 3D structure of SARS-CoV-2 (6VXX, closed state; 6VYB, closed state) to label the affected glycosylation sites (Fig. 5B-G). Among the sites affected by CEP on Spike protein, N30 and N122 are located at NTD region, N331 and N343 are in CTD/RBD region, and N1070 and N1094 are in S2 region²². As can be seen from the cartoon image (Figure 5G), the CEP-affected Spike protein glycosylation sites are located at both ends of the S protein. Four of the sites (including those located in the NTD and CTD/RBD regions) are at one end away from the virion, while the sites affected by the two sites affected locate in the S2 region at one end close to the virion. This suggests that one of the effects of CEP on viral S proteins might be to interfere with viral assembly and viral recognition or binding to the S protein receptors to exert its antiviral effects.

Multi-target drug combination is essential for the treatment of coronavirus infestation

Through the analysis of coronavirus GX_P2V-induced host cell glycoproteomics, we found that the viral effect on infected host cells is multifaceted. It involved in dysfunction of multiple proteins and pathways, including ECM-receptor interaction, extracellular matrix organization, response to wounding, platelet degranulation, and etc. (Figure 1 and Figure 2). Our data uncovered coronavirus GX_P2V-induced aberrant glycosylation of a variety of proteins on membrane (Figure 3). There are more than 75% of aberrant glycosylation of proteins after GX_P2V infection in correspondence to membrane-related subcellular structures, involving cell membrane, endoplasmic reticulum membrane, lysosomal membrane, Golgi apparatus membrane and nuclear membrane (Figure 3). These abnormal changes in glycoproteins caused by coronaviruses may play an important role in viral invasion of host cells, resulting in cellular dysfunction.

We found that the antiviral mechanism of CEP is also diverse, as it exerts its antiviral function by regulating the glycosylation of about 12 proteins which locate in endoplasmic reticulum, lysosomal, nucleus, cell membrane and ECM (Figure 3). However, at least 57 aberrant glycoproteins caused by coronavirus were not significantly improved by CEP treatment. Therefore, multi-target drug use in combination with CEP is essential for the treatment of coronavirus infection.

Taken together, the results highlight the differentially expressed glycoproteins targets and related pathways in coronavirus infection and CEP intervention. Generally, coronavirus GX_P2V-induced aberrant N-glycosylation alteration in infected cells are prevalent. CEP can reduce the N-glycosylation of proteins located on cell membrane surface, indicating that CEP may achieve the protective effect against virus invasion by interfering potential N-glycoprotein targets in the affected cells. Besides, CEP can regulate the glycosylation of viral proteins S, M, and N, revealing targets or pathways relevant for viral pathogenicity. The enhancement of glycan shield on the viral proteins S and M might weaken coronavirus' binding to its recognition targets. Nevertheless, the dramatic down-regulation of N-glycosylation witnessed at the N25 site in viral N protein caused need further study.

This study provides a landscape of N-glycoproteomic profiling using the established SARS-CoV-2 cellular model, suggesting that it is of great importance in therapeutic target screening for drug discovery. The

results indicated that coronavirus can cause alterations in the glycosylation of proteins at multiple levels in infected cells, whereas CEP can partially maintain GX_P2V-induced aberrant N-glycoprotein targets as well as partially regulate the glycosylation of viral proteins, suggesting that multi-target drug combination is essential for the treatment of coronavirus infestation so as to achieve optimal therapeutic results. These aberrant N-glycosylation of GX_P2V infestation might be potential targets for combination therapy.

Methods

Cell culture and coronavirus

Vero E6 cells (American Type Culture Collection, Manassas, VA, USA) (ATCC, No. 1586) were grown in high-glucose containing DMEM medium (Gibco) supplemented with 10% fetal bovine serum (FBS) in 37 °C incubator with 5% CO₂. SARS-CoV-2 related coronavirus GX_P2V/pangolin/2017/Guangxi was isolated from Vero E6 cells from a dead smuggled pangolin in 2017, and its complete genome has been submitted to GenBank²³. The compound Cepharranthine (CEP) was purchased from TOPscience (Shanghai, China).

Cepharranthine (CEP) and GX_P2V treatment

Vero E6 cells were cultured in T175 flask and treated with or without GX_P2V (MOI of 0.01) in presence or absence of CEP in triplicates. Group 1 (Vero): normal control; group 2 (Vero_C): CEP only treatment for 72 hours; group 3 (Vero_P): GX_P2V treatment for 72 hours; group 4 (Vero_C_P): GX_P2V and CEP GX_P2V treatment for 72 hours. The cyto toxicity of CEP to Vero E6 cells was measured by CellTiter blue according to the manufacturer's protocol (Promega, Catalog Number: PR-G8081; Madison, WI, USA). Cell samples were harvested according to the protocol for Glycoproteomics analysis from PTM-Biolab (Hangzhou, China), and stored in -80°C refrigerator for further analysis.

Protein extraction and trypsin digestion

Samples stored in -80°C refrigerator were thawed on ice and centrifuged at 12,000 g at 4°C for 10 min to remove cell debris. The supernatant was transferred to a new centrifuge tube for protein concentration determination using BCA kit.

Equal amount of each protein sample was enzymatically lysed with same volume of lysis buffer containing the appropriate amount of reference protein. The protein samples were mixed with five times of the sample volume and precipitated with pre-cooled acetone at -20°C for 2 h, followed by centrifugation at 4500 g for 5 min, the precipitate was collected and washed twice with pre-cooled acetone. The protein precipitate from each sample was dried, and then sonicated in buffer containing 200 mM Triethylammonium bicarbonate (TEAB). The digestion was performed by incubation with trypsin at a ratio of 1: 50 (enzyme: protein, m/m) with protein samples for overnight. Finally, the samples were desalted according to the C18 ZipTips instructions, and vacuum freeze-dried for HPLC analysis.

The enrichment of glycosylation modification of peptides

The peptide fragments were dissolved in 40 μ L of enrichment buffer (80% acetonitrile/1% trifluoroacetic acid) and transferred to a hydrophilic micro-column. The enrichment was completed by centrifugation at 4000 g for approximately 15 min in HILIC. The hydrophilic micro-column was washed for 3 times with enrichment buffer. The glycopeptides were eluted with 10% acetonitrile, then the eluate was collected and vacuum dried. After drying, the eluate was reconstituted in 50 μ L of 50 mM ammonium bicarbonate buffer dissolved in 50 μ L of hydrogen peroxide with 2 μ L of water. The glycopeptides were incubated with PNGase F glycosidase overnight at 37 °C. Finally, salt was removed according to the C18 ZipTips instructions and vacuum freeze-dried for liquid-liquid analysis.

Liquid Chromatography-Mass Spectrometry (LC-MS)

The peptides were dissolved in liquid chromatography mobile phase A (0.1% (v/v) formic acid aqueous solution) and then separated using NanoElute ultra-efficient liquid phase system. Mobile phase A is an aqueous solution containing 0.1% formic acid; mobile phase B is an acetonitrile solution containing 0.1% formic acid. Liquid phase gradient setting: 0-50 min, 2%~22%B; 50-52 min, 22%~35%B; 52-55 min, 35%~90%B; 55-60 min, 90%B, the flow rate is maintained at 450 nL/min.

The peptides were separated by an ultra-high-performance liquid phase system and injected into a capillary ion source. Then the peptide segments are analyzed by a TIMS-TOF Pro mass spectrometer. The ion source voltage was set to 1.6 kV and both the peptide parent ion and its secondary fragments were detected and analyzed using TOF. The secondary mass spectrometry scan range was set to 100-1700 m/z. The data acquisition mode used parallel cumulative serial fragmentation (PASEF) mode. One primary mass spectrometry acquisition followed by 10 PASEF mode acquisitions of secondary spectra with parent ion charge numbers in the range 0-5. The dynamic exclusion time of the tandem mass spectrometry scan was set to 30 seconds to avoid repeated scans of the parent ion.

Database Search

Retrieval parameter settings: the database was FA105LPNg_Chlorocebus_sabaeus_60711_Ensembl_GX_P2V_protein_TX_combine_20200413 (19267 sequences), an anti-library was added to calculate the false positive rate (FDR) caused by random matching, and a common pollution library was added to the database to eliminate the contamination protein in the identification results Impact; the digestion method was set to Trypsin/P; the number of missed cleavage sites was set to 2; the minimum length of the peptide was set to 7 amino acid residues; the maximum modification number of the peptide was set to 5.

The mass error tolerance of the primary precursor ions in First search and Main search was respectively set to 20 ppm and 20 ppm, and the mass error tolerance of the secondary fragment ions was 0.02 Da. The cysteine alkylation was set as a fixed modification, and the variable modification was the oxidation

of methionine, the acetylation of the N-terminus of the protein, and the deamination of asparagine (180). The FDR for protein identification and PSM identification was set to 1%.

Bioinformatics analyses

Glycoproteomic data of GX_P2V infection group and 14 glycosylation sites in 12 glycoprotein targets of CEP treatment in presence of GX_P2V I was analyzed by Metascape web-based platform²⁴. Pathway and process enrichment analysis were carried out with the following ontology sources: KEGG Pathway, GO Biological Processes, Reactome Gene Sets, Canonical Pathways and CORUM. PPI networks were constructed by string (STRING; <http://string-db.org>. version 11.0)²⁵ and Cytoscape (version 3.7.1)²⁶, the most relevant modules in the PPI networks were identified using MCODE²⁷, and the network adjusted by the Files Layout algorithm. Reactome analysis tool (<http://reactome.org>) was used to identify the enriched pathways of glycosylation sites on glycoproteins remains at normal state in response to GX_P2V treatment in presence of CEP. ($p \leq 0.05$) (<https://reactome.org/PathwayBrowser/#TOOL=AT>)²⁸. Uniprot was used to find the subcellular localization of proteins (<https://www.uniprot.org/>)¹³. One-way ANOVA followed by Dunnett's multiple comparisons test was performed using GraphPad Prism version 8.0.0 for Windows, GraphPad Software, San Diego, California USA, www.graphpad.com.

Data Availability

All data generated or analyzed during this study are presented in this paper or in the Supplementary Information. The MS proteomics data have been deposited in the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD021297.

Declarations

Author contributions

W. An, H. Fan and Y. Tong conceived, designed and supervised this study. F. Tian and J. Li analyzed the data, carried out literature search and wrote the initial manuscript. J. Li prepared most of the figures and tables. Y. Wang provided assistance for data analyses and literature search. W. An, H. Fan and S. Wang edited, corrected and proofread the full contents of the paper. All authors read and approved the final manuscript.

Acknowledgments

This research was supported by a grant from The National Key Research and Development Program of China (2018YFA0903000), the National Natural Science Foundation of China (81672001), Fundamental Research Funds for Central Universities (No. BUCTRC201917, No. BUCTZY2022) and China MoST Emergency Project on COVID-19 (2020YFC0840800).

Competing interests

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

References

1. Hon KL, *et al.* Coronavirus disease 2019 (COVID-19): latest developments in potential treatments. *Drugs in context* **9**, (2020).
2. Ma R, Peng X, Xu Y, Duan DY. [Advances in salivary protein glycosylation and its relationship with systemic and oral diseases]. *Hua xi kou qiang yi xue za zhi = Huaxi kouqiang yixue zazhi = West China journal of stomatology* **36**, 336-341 (2018).
3. Dall'Olio F, Vanhooren V, Chen CC, Slagboom PE, Wuhrer M, Franceschi C. N-glycomic biomarkers of biological aging and longevity: a link with inflammaging. *Ageing research reviews* **12**, 685-698 (2013).
4. Gornik O, Wagner J, Pucić M, Knezević A, Redzic I, Lauc G. Stability of N-glycan profiles in human plasma. *Glycobiology* **19**, 1547-1553 (2009).
5. Clerc F, Reiding KR, Jansen BC, Kammeijer GS, Bondt A, Wuhrer M. Human plasma protein N-glycosylation. *Glycoconjugate journal* **33**, 309-343 (2016).
6. Robson B. Bioinformatics studies on a function of the SARS-CoV-2 spike glycoprotein as the binding of host sialic acid glycans. *Computers in biology and medicine* **122**, 103849 (2020).
7. Fan HH, *et al.* Repurposing of clinically approved drugs for treatment of coronavirus disease 2019 in a 2019-novel coronavirus-related coronavirus model. *Chinese medical journal* **133**, 1051-1056 (2020).
8. Spiro RG. Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology* **12**, 43r-56r (2002).
9. Rudd PM, Elliott T, Cresswell P, Wilson IA, Dwek RA. Glycosylation and the immune system. *Science (New York, NY)* **291**, 2370-2376 (2001).
10. Raghunathan R, Sethi MK, Klein JA, Zaia J. Proteomics, Glycomics, and Glycoproteomics of Matrisome Molecules. *Molecular & cellular proteomics : MCP* **18**, 2138-2148 (2019).
11. Patarroyo M, Tryggvason K, Virtanen I. Laminin isoforms in tumor invasion, angiogenesis and metastasis. *Seminars in cancer biology* **12**, 197-207 (2002).
12. Ye G, *et al.* Lamc1 promotes the Warburg effect in hepatocellular carcinoma cells by regulating PKM2 expression through AKT pathway. *Cancer biology & therapy* **20**, 711-719 (2019).
13. UniProt C. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res* **47**, D506-D515 (2019).
14. Sheikh H, Yarwood H, Ashworth A, Isacke CM. Endo180, an endocytic recycling glycoprotein related to the macrophage mannose receptor is expressed on fibroblasts, endothelial cells and macrophages and functions as a lectin receptor. *Journal of cell science* **113 (Pt 6)**, 1021-1032 (2000).

15. Wienke D, MacFadyen JR, Isacke CM. Identification and characterization of the endocytic transmembrane glycoprotein Endo180 as a novel collagen receptor. *Molecular biology of the cell* **14**, 3592-3604 (2003).
16. Zhou X, *et al.* Impaired prosaposin lysosomal trafficking in frontotemporal lobar degeneration due to progranulin mutations. *Nature communications* **8**, 15277 (2017).
17. Beel S, *et al.* Progranulin functions as a cathepsin D chaperone to stimulate axonal outgrowth in vivo. *Human molecular genetics* **26**, 2850-2863 (2017).
18. Watanabe Y, Allen JD, Wrapp D, McLellan JS, Crispin M. Site-specific analysis of the SARS-CoV-2 glycan shield. *bioRxiv*, (2020).
19. Watanabe Y, *et al.* Vulnerabilities in coronavirus glycan shields despite extensive glycosylation. *Nature communications* **11**, 2688 (2020).
20. Watanabe Y, Allen JD, Wrapp D, McLellan JS, Crispin M. Site-specific glycan analysis of the SARS-CoV-2 spike. *Science (New York, NY)*, (2020).
21. Guruprasad L. Evolutionary relationships and sequence-structure determinants in human SARS coronavirus-2 spike proteins for host receptor recognition. *Proteins*, (2020).
22. Wang Q, *et al.* Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell* **181**, 894-904.e899 (2020).
23. Szklarczyk D, *et al.* STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* **47**, D607-d613 (2019).
24. Lam TT, *et al.* Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins. *Nature* **583**, 282-285 (2020).
25. Zhou Y, *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nature communications* **10**, 1523 (2019).
26. Doncheva NT, Morris JH, Gorodkin J, Jensen LJ. Cytoscape StringApp: Network Analysis and Visualization of Proteomics Data. *J Proteome Res* **18**, 623-632 (2019).
27. Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. *BMC bioinformatics* **4**, 2 (2003).
28. Haw R, Hermjakob H, D'Eustachio P, Stein L. Reactome pathway analysis to enrich biological discovery in proteomics data sets. *Proteomics* **11**, 3598-3613 (2011).

Figures

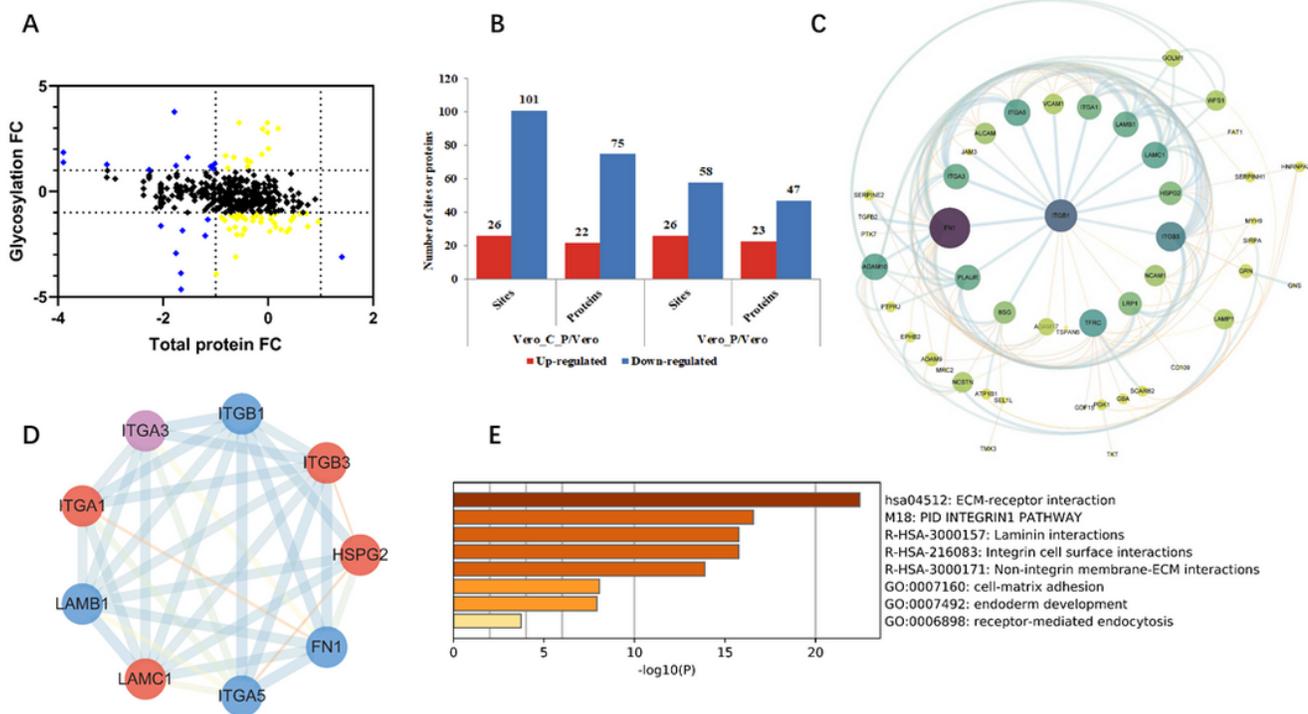


Figure 1

Effects of GX_P2V infection on the aberrant glycosylation of proteins in Vero E6 cells. (A) Global presentation of proteomic and respective glycosylation occupancy. Yellow dots represent proteins which have not changed significantly in global proteomics data but the peptides with glycosylation have differentially expressed over 2 folds, and blue dots represent proteins which are differentially expressed both in the levels of global protein and glycosylation. (B) Effects of GX_P2V treatment on glycosylation (sites or proteins) in presence or absence of CEP. In comparison with Vero E6 group (control), GX_P2V treatment can up-regulate 26 glycosylation sites in 23 proteins and down-regulate 58 glycosylation sites in 47 proteins, whereas in presence of CEP, GX_P2V treatment can induce 26 glycosylation sites up-regulated in 22 proteins and 110 glycosylation sites down-regulated in 75 proteins in Vero E6 cells. Fold change (Vero_C_P/Vero or Vero_P/Vero) ≥ 2 or ≤ 0.5 was considered as up-regulation or down-regulation, respectively. (C) Protein-protein interaction (PPI) network and the interaction modules of key proteins in response to GX_P2V treatment. The PPI network of 70 differentially expressed glycoproteins ($Fc \geq 2$ or $Fc \leq 0.5$) in GX_P2V treatment was produced using Cytoscape. (D) The key interaction module was obtained from the PPI network by Cytoscape with MCODE plugin using differentially expressed glycosproteins in GX_P2V-treated Vero E6 cells. Proteins marked in red color represent those proteins containing up-regulated glycosylation sites, the blue ones indicate proteins containing down-regulated glycosylation sites, and the purple color labelled protein contains both up- and down-regulated glycosylation sites. (E) Functional enrichment analysis of the key modules induced by GX_P2V treatment was obtained from Metascape.

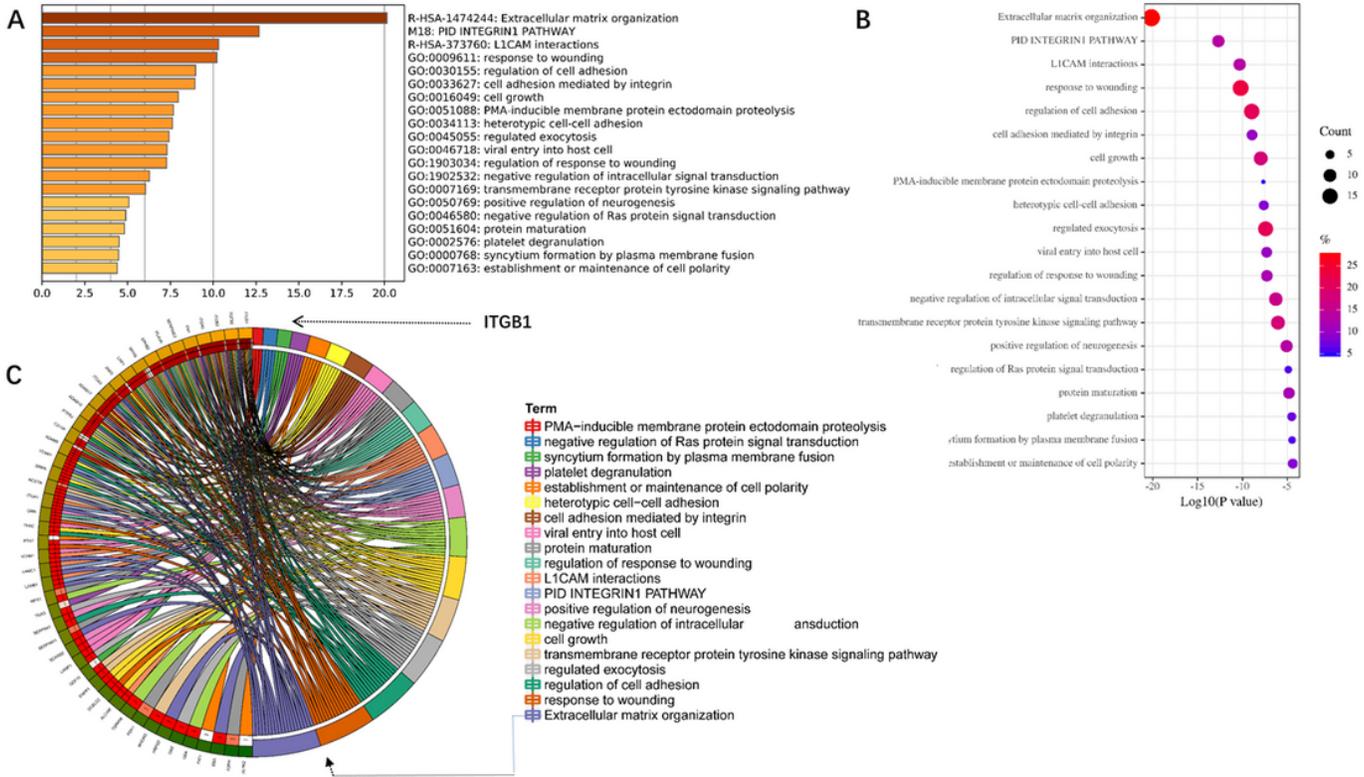


Figure 2

Functional enrichment clustering analysis of differentially expressed glycoproteins in GX_P2V -treated Vero E6 cells. (A) Functional enrichment analysis of proteins with aberrant glycosylation induced by GX_P2V infection was constructed by Metascape. (B) Enriched bubble diagram of proteins with aberrant glycosylation after GX_P2V treatment was drawn by SangerBox. (C) Circle diagram of protein functional enrichment analysis showing the relationship of aberrant glycosylation sites and the involved functions. The inner circle on the left side represents the significance p value of the gene corresponding pathway, and the right represents the corresponding term. The denseness of the left semicircular line represents the number of pathways associated with the protein. The density of the right semicircle line indicates how many proteins are enriched in that pathway. The figure shows that ITGB1 is involved in the most relevant pathways and the extracellular matrix is enriched for the most genes.

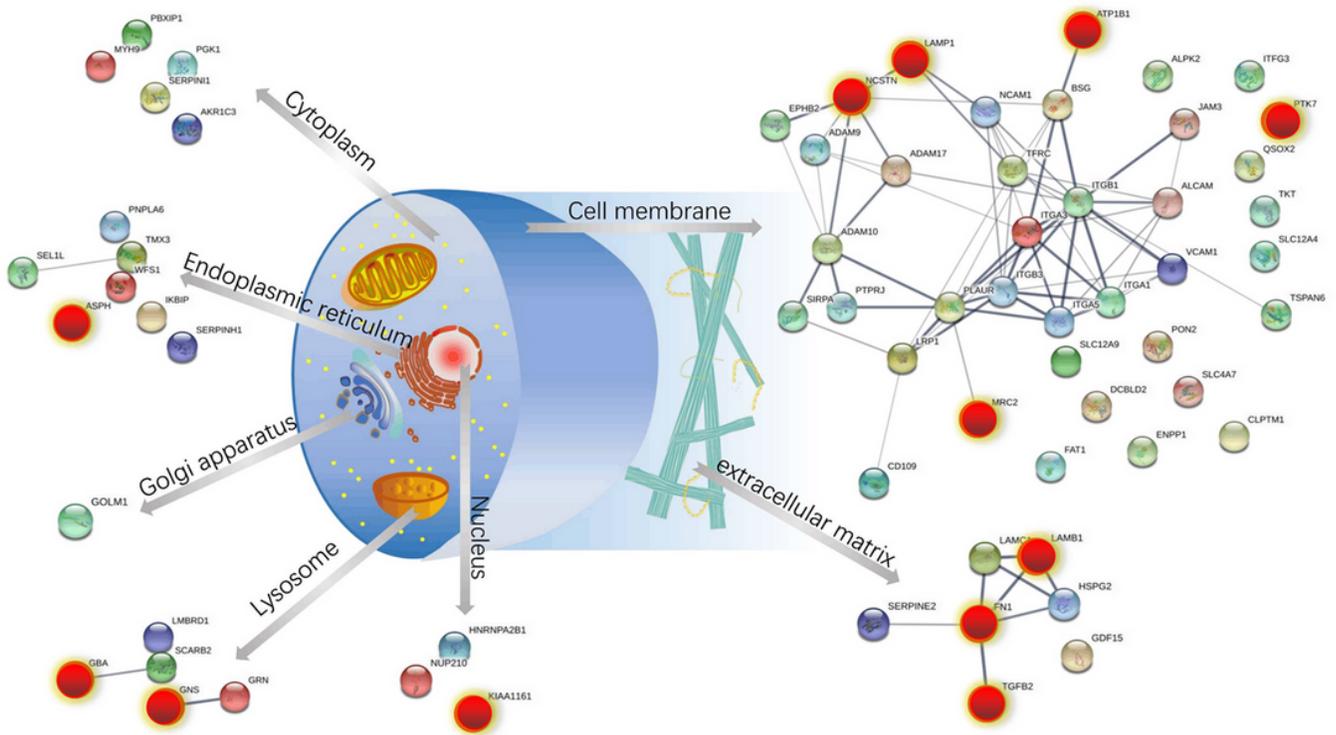


Figure 3

Schematic diagram of subcellular location and PPI of differentially expressed glycoproteins in GX_P2V treated Vero E6 cells as well as the intervention targets of CEP. The subcellular location of aberrant glycoproteins was determined using UniProt. The PPI network maps of these dysregulated glycoproteins were generated using STRING. The red color highlights CEP-regulated proteins. Upon GX_P2V treatment in presence of CEP, 12 out of 69 glycoproteins dysregulated by coronavirus GX_P2V remains in normal glycosylation state, suggesting that these proteins and glycosylation sites might be the targets of CEP.

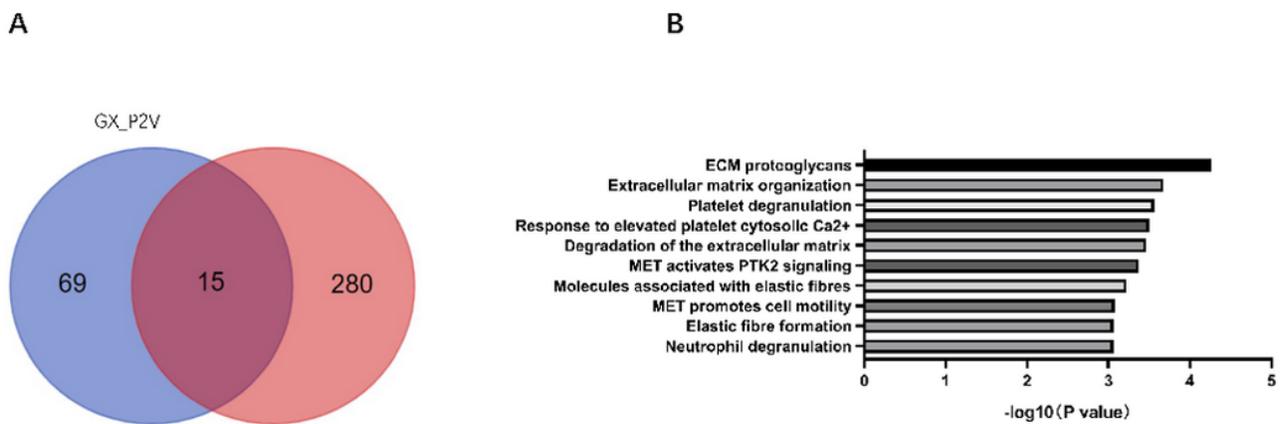


Figure 4

Effects of CEP treatment on GX_P2V-induced aberrant glycosylation of proteins in Vero E6 cells. (A) Venn diagram shows the effect of CEP on the aberrant glycosylation of proteins induced by GX_P2V. The number in purple circle indicates the glycosylation sites those are differentially expressed (with Vero_P/ Vero Fc ≥ 2) in response to GX_P2V treatment (GX_P2V). The number in light red circle shows the glycosylation sites are in normal state (with Vero_C_P/ Vero, $0.667 \geq Fc \leq 1.5$) in presence of CEP (CEP+GX_P2V). The number in the overlap of the two circle shows that there are 14 site (on 12 proteins) out of 84 sites (on 69 proteins) GX_P2V-induced aberrant glycosylation sites in normal state in presence of CEP. (B) Reactome Pathway Analysis of glycosylation sites on 12 glycoproteins affected by CEP.

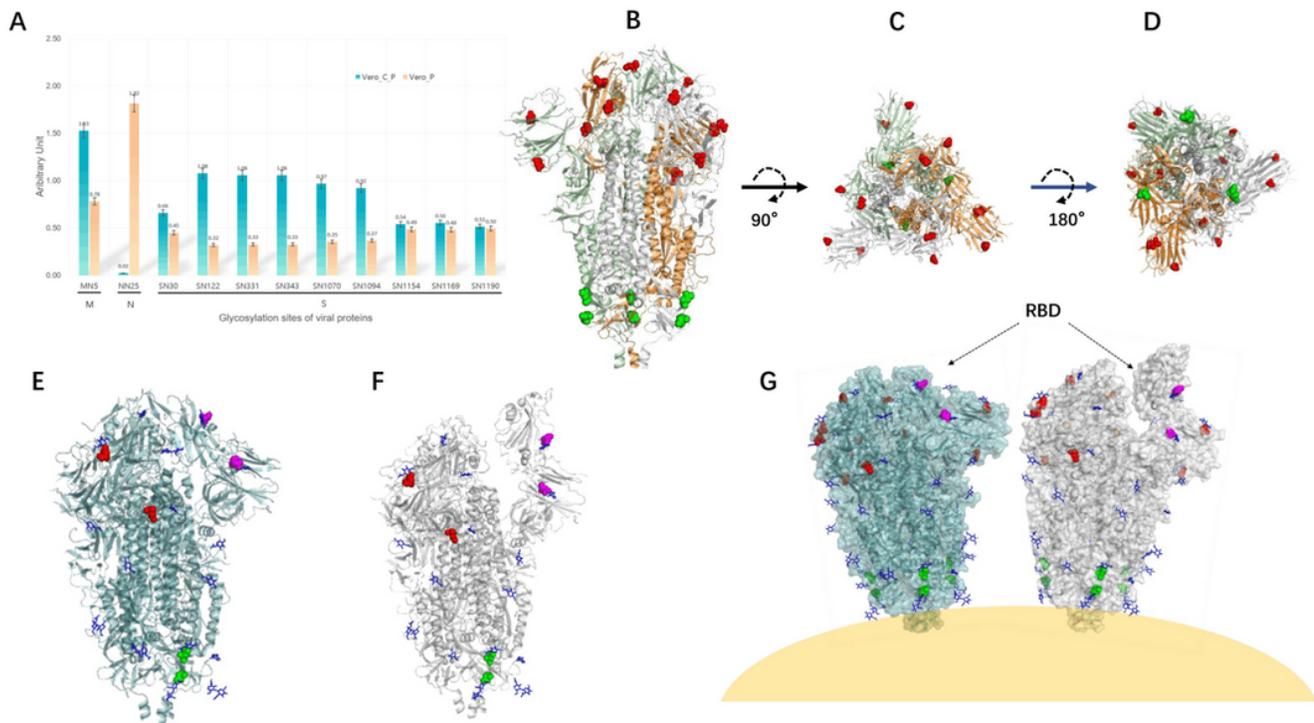


Figure 5

Effect of CEP on the N-Glycosylation of GX_P2V viral Proteins. (A) Expression levels of N-glycosylation of viral proteins at indicated sites in presence or absence of CEP. The error bars represent the standard standard deviation; (B-D) indicate the side, top and bottom views of viral Spike protein (in closed state, PDB: 6VXX), respectively. All the N-glycosylation sites on the three chains of S protein affected by CEP are labelled. The N-glycosylation sites close to virus body are labelled in green, and those sites located at far-end of virus body are labelled in red. (E-F) Side views of viral Spike protein (the N-glycosylation sites affected by CEP are labelled on B chain (Figure 6E, PDB: 6VXX, closed state; Figure 6F, PDB: 6VYB, open state). Blue, NAG, N-linked glycan to the B chain of viral Spike protein; Magenta, the N-glycosylation sites located on the RBD of viral S protein; Green, the N-glycosylation sites affected by CEP on B chain of S protein close to virus body. Red, the CEP-affected N-glycosylation sites located on N-terminus of B chain. (G) Cartoon display of CEP-affected N-glycosylation sites on Spike protein on the coronavirus surface. Note: Since the 3D structure of GX_P2V viral S proteins is not available, considering the high sequence

similarity between SARS-CoV-2 and GX_P2V, we adopted the 3D structure of SARS-CoV-2 (6VXX, closed state; 6VYB closed state) to draw the pictures

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementary.docx](#)