

Identification of 22 N-glycosites on Spike glycoprotein of SARS-CoV-2 and accessible surface glycopeptide motifs: implications on vaccination and antibody therapeutics

Dapeng Zhou^{1*#}, Xiaoxu Tian^{2#}, Ruibing Qi³, Chao Peng^{2*}, Wen Zhang^{4,5 *}

¹*Tongji University School of Medicine, Shanghai, 200092, China;*

²*National Facility for Protein Science in Shanghai, Zhangjiang Lab, Shanghai Advanced Research Institute, Chinese Academy of Science, Shanghai 201210, China;*

³*Innovation Team of Small Animal Infectious Disease, Shanghai Veterinary Research Institute, Chinese Academy of Agricultural Science, Shanghai 200241, China*

⁴*Fudan University Pudong Medical Center, Institutes of Biomedical Sciences;* ⁵*Department of Systems Biology for Medicine, Shanghai Medical College, Fudan University, 200032, China*

*Corresponding authors:

dapengzhoulab@tongji.edu.cn

pengchao@sari.ac.cn

wenz@fudan.edu.cn

Contributed equally to this work

Supplemental Online Materials

Supplemental Table 1: Peptide and Glycopeptide mixtures of trypsin and chymotrypsin-digested SARS-CoV-2 Spike protein identified by LC-MS. Red bars indicate molecular ions identified by the software in pGlyco2.0. A total of 22 N-glyco-sites were identified and further confirmed by MS-MS analysis (Supplemental Figure 1).

Supplemental Table 2: List of trypsin and chymotrypsin-digested glycopeptides of SARS-CoV-2 Spike protein identified by LC-MS. The majority of N-glycans are high mannose type.

Supplemental Table 3: List of monoclonal antibodies for Spike protein of SARS-CoV-1

Supplemental Figure 1: MS-MS spectrum of glycan moieties and b/y ions for glycopeptides of 22 N-glycosites. MS² spectrum was automatically annotated and displayed by the software tool gLabel embedded in pGlyco2.0. “J” indicates the N-glycosylation site “N”. Colored circles indicate the number of sugars in attached glycan: green circle, hexose (H); blue square, N-acetylglucosamine (N); red triangle, fucose (F). Colored peaks in MS² spectrum include: green peaks representing the fragment ions of a glycan moiety; blue peaks representing a diagnostic glycan ion; red peaks representing the Y ions from glycan fragmentation; and yellow/cyan peaks representing the b/y ions from peptide backbone fragmentation. Mass deviations of the annotated peaks are shown in the lower box.

Supplemental Figure 2: Structure-based alignment of SARS-CoV-2 (2019-nCoV) and SARS-CoV-1 Spike proteins. The sequences are directly extracted from PDB 5X58 and 2019-nCoV homology model, and the sequence alignment was based on above two structures by ENDscript and ESPRIPT with default settings (<http://espruit.ibcp.fr/ESPrut/ENDscript/index.php>).

Supplemental Figure 3: Accessible surface area profiling of Spike proteins of SARS-CoV-2 (2019-nCoV) and SARS-CoV-1. A) The epitopes predicted on the S protein structure for SARS-CoV, Epi (yellow) denotes the epitopes screened by simple ASA profiling (the same for nCoV), and EpiS (red) denotes the epitopes were calculated by excluding the glycosylation sites

and the glyco-interacting amino acids; B) The epitopes predicted for nCoV. The values of Y axis means nm² of ASA.

Supplemental Figure 4: Connecting region (CR) of SARS-CoV-2 (2019-nCoV) and SARS-CoV-1 Spike proteins.

Supplemental Figure 5: Furin recognition site of SARS-CoV-2 (2019-nCoV) Spike protein.