

## Binding of the SARS-CoV-2 Spike Protein to Glycans

Wei Hao<sup>1,†</sup>, Bo Ma<sup>2,†</sup>, Ziheng Li<sup>1</sup>, Xiaoyu Wang<sup>2</sup>, Xiaopan Gao<sup>1</sup>, Yaohao Li<sup>2,3</sup>, Bo Qin<sup>1</sup>, Shiyang Shang<sup>4</sup>, Sheng Cui<sup>1,\*</sup>, Zhongping Tan<sup>2,\*</sup>

<sup>1</sup> NHC Key Laboratory of Systems Biology of Pathogens, Institute of Pathogen Biology and Center for AIDS Research, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100730, China.

<sup>2</sup> State Key Laboratory of Bioactive Substance and Function of Natural Medicines, Institute of Materia Medica, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100050, China

<sup>3</sup> Department of Chemistry and Biochemistry and BioFrontiers Institute, University of Colorado, Boulder CO 80303, United States

<sup>4</sup> School of Pharmaceutical Sciences, Tsinghua University, Beijing 100084, China

### \* Correspondence:

Corresponding Author

cui.sheng@ipb.pumc.edu.cn; zhongping.tan@imm.pumc.edu.cn

**Keywords:** Heparan sulfate, Glycan microarray, Surface plasmon resonance, Sulfation, Sialic acid

### Abstract

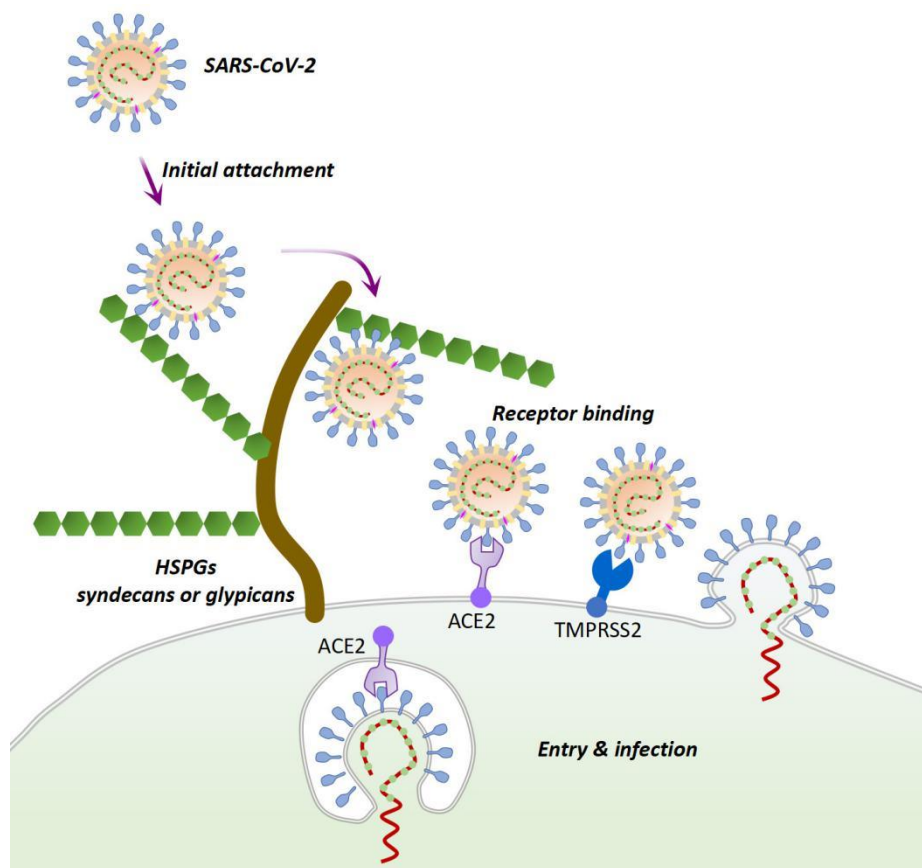
The pandemic of SARS-CoV-2 has caused a high number of deaths in the world. To combat it, it is necessary to develop a better understanding of how the virus infects host cells. Infection normally starts with the attachment of the virus to cell-surface glycans like heparan sulfate (HS) and sialic acid-containing oligosaccharides. In this study, we examined and compared the binding of the subunits and spike (S) proteins of SARS-CoV-2 and SARS-CoV, MERS-CoV to these glycans. Our results revealed that the S proteins and subunits can bind to HS in a sulfation-dependent manner, the length of HS is not a critical factor for the binding, and no binding with sialic acid residues was detected. Overall, this work suggests that HS binding may be a general mechanism for the attachment of these coronaviruses to host cells, and supports the potential importance of HS in infection and in the development of antiviral agents against these viruses.

## Introduction

The 2019 novel coronavirus (CoV) is the seventh human coronavirus.<sup>1</sup> It is a deadly virus that is affecting the whole world in an unprecedented way. The global impact of the coronavirus disease 2019 (COVID-19) pandemic is far beyond that of two other major coronavirus outbreaks in the past 20 years, the severe acute respiratory disease (SARS) in 2003<sup>2,3</sup> and the Middle East respiratory disease (MERS) in 2012.<sup>4</sup> Given that all three highly pathogenic CoVs were originated from bats and a large number of closely related CoVs are present in bats, future outbreak of this type of zoonotic virus remains possible. In order to avoid facing a similar pandemic in the future, it is necessary to develop a better understanding of these CoVs, especially with regard to effective ways that can help control the current COVID-19 pandemic and prevent the second wave of outbreaks.

Studies showed that the genome of the SARS-CoV-2 has about 80% nucleotide identity with that of SARS-CoV.<sup>5</sup> The major differences are found in the regions encoding the structural proteins (envelope E, membrane M, nucleocapsid N, and spike S) and accessory proteins (ORF3a/3b, 6, 7a/7b, 8 and 10), not the nonstructural proteins (nsp1 to nsp16). Based on this genetic similarity, the 2019 novel CoV was named by the International Committee on Taxonomy of Viruses (ICTV) as the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2).<sup>6</sup> SARS-CoV-2 is more genetically distant from MERS-CoV and shares only about 57% genome homology with MERS-CoV.<sup>5,7</sup> These similarities and differences are reflected in the viral binding receptors to cell surface receptors. SARS-CoV-2 and SARS-CoV use the same receptor angiotensin-converting enzyme 2 (ACE2) for entry into target cells,<sup>8-10</sup> whereas MERS-CoV uses dipeptidyl peptidase 4 (DPP4, also known as CD26) as the primary receptor.<sup>11</sup> However, there are many important findings that cannot be explained by their genetic relations. For example, it was found that SARS-CoV-2 has a relatively lower mortality rate (2%) than SARS-CoV (15%) and MERS-CoV (35%), but is more transmissible among humans.<sup>12-14</sup> These findings suggest that more investigations other than genetic analysis need to be carried out to improve the understanding of SARS-CoV-2 infection.

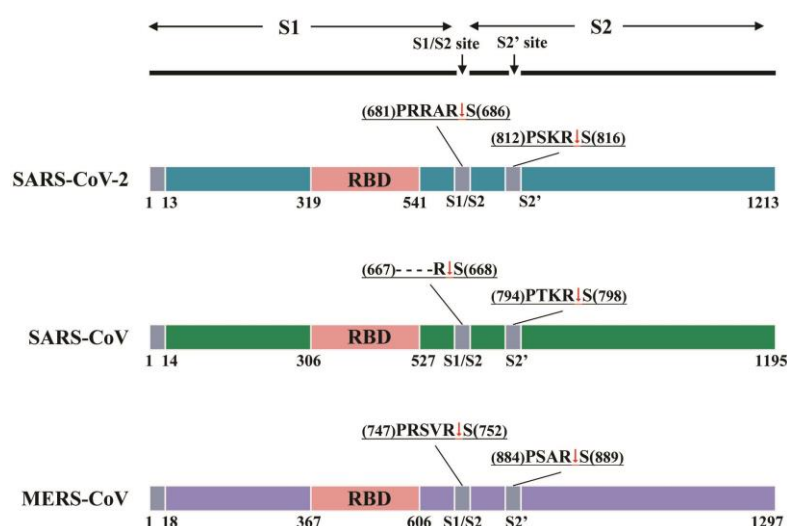
In order to efficiently infect host cells, SARS-CoV-2 must bind with cell surface molecules in the lungs and other organs to mediate viral attachment and entry into host cells. Previous studies of many other viruses suggested that SARS-CoV-2 S protein may use other molecules on host cell surface as attachment factors to facilitate binding to the high-affinity receptor ACE2.<sup>15,16</sup> Examples of such molecules include glycosaminoglycans (GAGs) and sialic acid-containing oligosaccharides.<sup>17-20</sup> GAGs are primarily localized at the outer surface of cells. Such a location makes them particularly suitable for acting as attachment factors to recruit viruses to cell surfaces.<sup>21,22</sup> HS is one of the most prevalent types of GAGs in mammals. It is a linear and sulfated polysaccharide that is abundantly expressed on the surface of almost all cell types and in the extracellular matrix.<sup>23,24</sup> The HS chains are mostly covalently linked as side chains to core proteins to form HS proteoglycans (HSPGs) (Fig. 1).



**Figure 1.** A possible mechanism for SARS-CoV-2 entry and infection. At the early stage of the infection process, SARS-CoV-2 may first interact with the HSPGs on the surface of susceptible cells using the S protein protruding from the virus particle. This initial attachment may promote the subsequent binding of the virus to the high-affinity entry receptor ACE2. The serine protease TMPRSS2 on host cell surface and other host cell proteases may assist in viral entry by cleaving the S protein at the S1/S2 and/or at the S2' sites.

HS is synthesized in the Golgi apparatus by many different enzymes. During and after its assembly, HS undergoes extensive series of modifications including sulfation, acetylation and epimerization, which leads to glycan structures with high heterogeneity in length, sulfation, and glucuronate/iduronate ratio.<sup>25</sup> Considerable variation in the sulfation pattern and degree of HS was noted in different species, organs, tissues, and even at different ages and disease stages.<sup>23,24,26</sup> The sequence and sulfation pattern of HS has been shown to be able to regulate the binding of many viruses to host cells during infection.<sup>27,28</sup> A similar trend was also observed for sialylation patterns of cell surface oligosaccharides.<sup>17,29,30</sup> These findings implicate that a possible relationship may exist between the distribution of different types of HS/sialylated glycans and the viral tropism.<sup>31-33</sup> A better understanding of their relationship can potentially contribute to the design of new antiviral strategies. Although currently the data on the viral tropism of SARS-CoV-2 are limited, results from recent studies suggested that its tropism may not be correlated with the ACE2 expression.<sup>34,35</sup> Some other factors, such as proteases and glycans, may be the determinant of cellular susceptibility to the infection with this virus.<sup>36</sup> A recent study suggested that HS may bind to the receptor binding domain

(RBD, the C-terminal region of the S1 subunit, Fig. 2) of the SARS-CoV-2 spike protein and change its conformation.<sup>37</sup> The intriguing possibility that variation in HS and sialic acid characteristics could impact the tropism of viruses prompts us to investigate the binding of SARS-CoV-2 toward a series of HS and sialic acid containing oligosaccharides.<sup>38</sup>



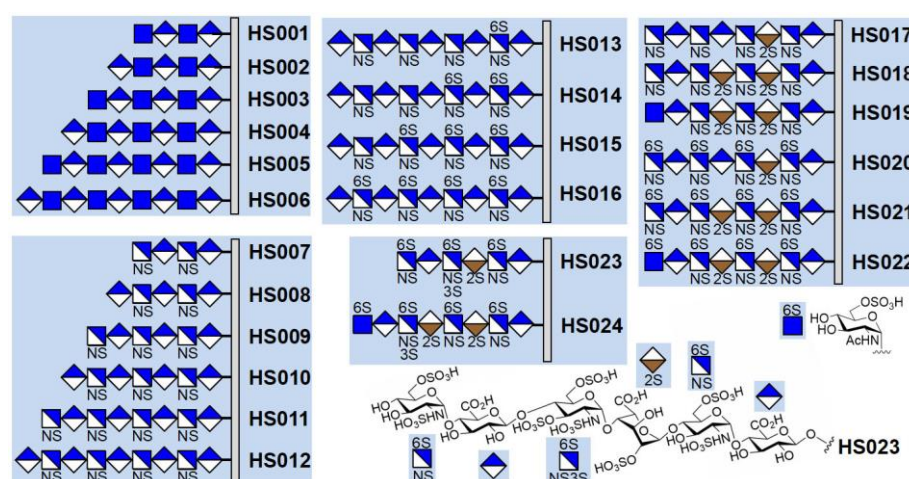
**Figure 2.** Schematic diagram showing the domain organization of the S proteins of SARS-CoV-2, SARS-CoV and MERS-CoV. The RBD is shown in pink color. The sequences of the S1/S2 and S2' cleavage sites are given on top of each diagram. The red arrows indicate the cleave points.

In this study, we systematically examined and compared the binding of the SARS-CoV-2 S protein subunits, full-length molecule and its trimer to different HS using microarray experiments (Fig. 2). Our results suggested that all the tested protein molecules were capable of binding HS oligomers and had similar binding preferences, with higher affinity toward HS forms with higher degrees of sulfation. The binding was also shown to be positively related to the level of 6-O-sulfation. The HS chain length seemed to be not critical. A long heparin molecule (a highly sulfated HS) and a synthetic heparin pentasaccharide (Fondaparinux sodium) were demonstrated to have similar binding affinity to these protein molecules. Moreover, our study suggested that the SARS-CoV-2 S protein might not be able to bind sialic acid residues, or the binding might be too weak to be detected by the microarray technology, and the S proteins of SARS-CoV and MERS-CoV have similar HS and sialic acid binding properties as those of the SARS-CoV-2 S protein. Overall, our study laid a foundation for future studies to explore whether the binding specificity to HS can serve as an important contributor to the viral tropism of SARS-CoV-2 and to explore the possibility of exploiting HS for therapeutic strategies.

## Results

**Binding of the subunits and S proteins of SARS-CoV-2, SARS-CoV, and MERS-CoV to a HS microarray.** An early study has suggested that the RBD of SARS-CoV-

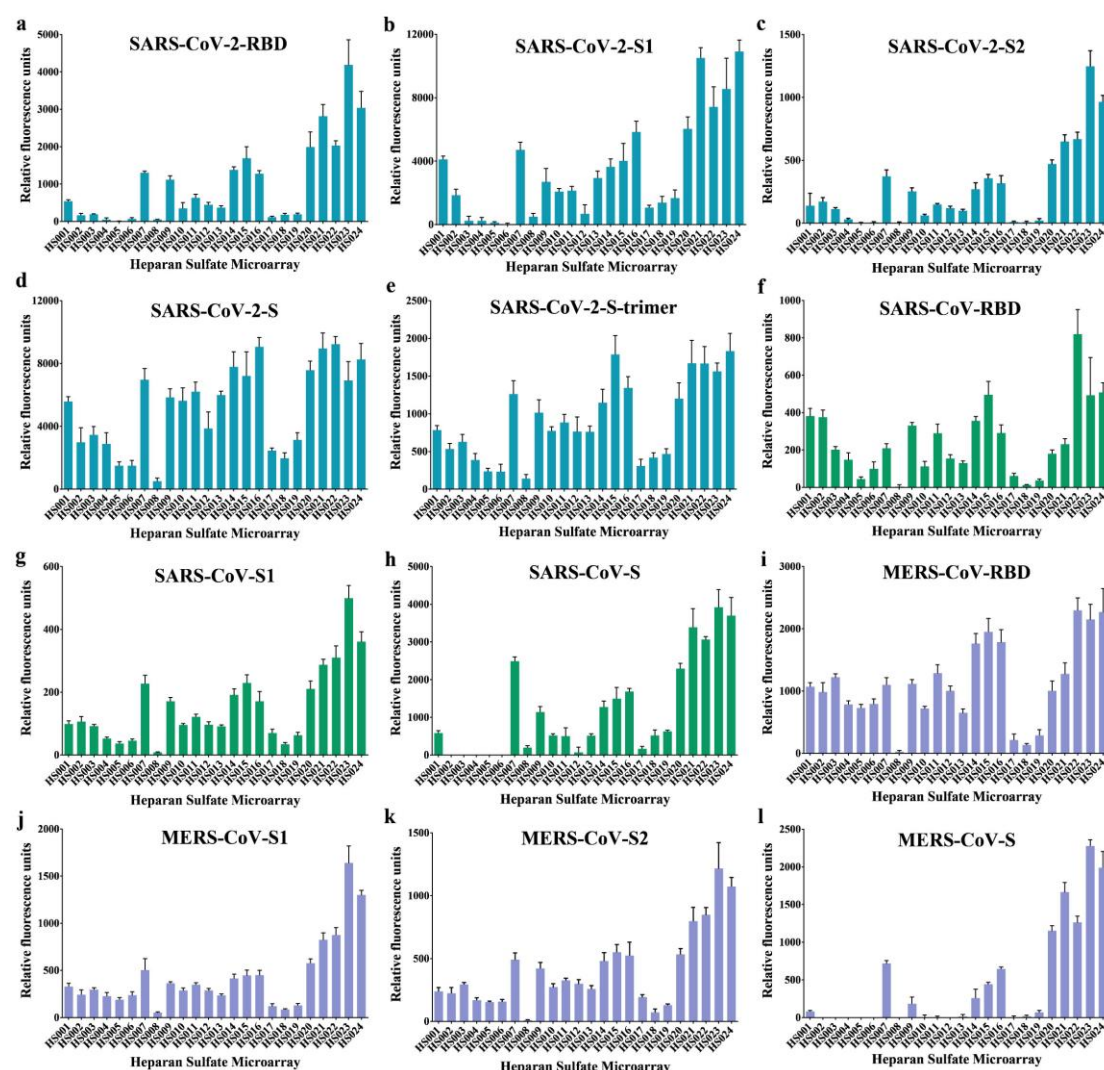
2 S protein might bind heparin.<sup>9</sup> In order to determine if there is any preference of the SARS-CoV-2 S protein for particular HS structures, we first investigated the binding of the RBD (Here termed as SARS-CoV-2-RBD) to a HS microarray containing 24 synthetic heparan sulfate oligosaccharides. These oligosaccharides have systematic differences in their length, monosaccharide composition, and sulfation pattern (Fig. 3). The microarray experiment was performed using a previously established standard protocol.<sup>29,39,40</sup> Briefly, the proteins were labeled with Cy3 fluorescent dye and incubated with the microarray at different concentrations. After washing away the unbound HS molecules, a highly sensitive fluorescence method was used to detect the binding of the SARS-CoV-2-RBD to HS. Under the experimental conditions, the binding can be detected at concentrations higher than 0.5  $\mu\text{g/ml}$ . Increasing the concentration of the protein was not found to produce noticeable changes in binding (Supporting Information, Fig. S1).



**Figure 3:** The numbering and structures of the HS oligosaccharides on the microarray. Each HS chain is covalently attached to the microarray slide via the reducing end. The HS oligosaccharides can be roughly divided into five groups according to their differences in the monosaccharide compositions and sulfation patterns.

Quantification of fluorescence revealed that the SARS-CoV-2-RBD is able to bind to almost half of the molecules on the microarray, and not surprisingly, the binding is strongly affected by the sulfation level, which is a trend that has been previously noted for many HS-binding proteins.<sup>41-43</sup> As shown in Figure 4a, the HS oligosaccharides with higher sulfation degree, **HS020-HS024** (the number of sulfate groups per monosaccharide unit  $>1.00$ ), exhibit higher fluorescence intensity. The highest fluorescence intensity is observed for **HS023** (1.35 sulfate groups per monosaccharide), which is followed by those of **HS021** and **HS024** (1.25 sulfate groups per monosaccharide) and those of **HS020** and **HS022** (1.15 sulfate groups per monosaccharide). Binding of SARS-CoV-2-RBD to HS seems to not be affected by the monosaccharide composition (compare **HS020** with **HS022**, and **HS021** with **HS024**).





**Figure 4.** The contributions of different HS characteristics to the binding to (a) SARS-CoV-2-RBD at the concentration of 2.5  $\mu\text{g/ml}$ , (b) SARS-CoV-2-S1 at the concentration of 2  $\mu\text{g/ml}$ , (c) SARS-CoV-2-S2 at the concentration of 2  $\mu\text{g/ml}$ , (d) SARS-CoV-2-S at the concentration of 2  $\mu\text{g/ml}$ , (e) SARS-CoV-2-trimer at the concentration of 2  $\mu\text{g/ml}$ , (f) SARS-CoV-RBD at the concentration of 2  $\mu\text{g/ml}$ , (g) SARS-CoV-S1 at the concentration of 2  $\mu\text{g/ml}$ , (h) SARS-CoV-S at the concentration of 2  $\mu\text{g/ml}$ , (i) MERS-CoV-RBD at the concentration of 4  $\mu\text{g/ml}$ , (j) MERS-CoV-S1 at the concentration of 4  $\mu\text{g/ml}$ , (k) MERS-CoV-S2 at the concentration of 2  $\mu\text{g/ml}$ , (l) MERS-CoV-S at the concentration of 2  $\mu\text{g/ml}$ . The fluorescence intensity was measured at an excitation wavelength of 532 nm. All error bars are standard deviation of more than three replicates.

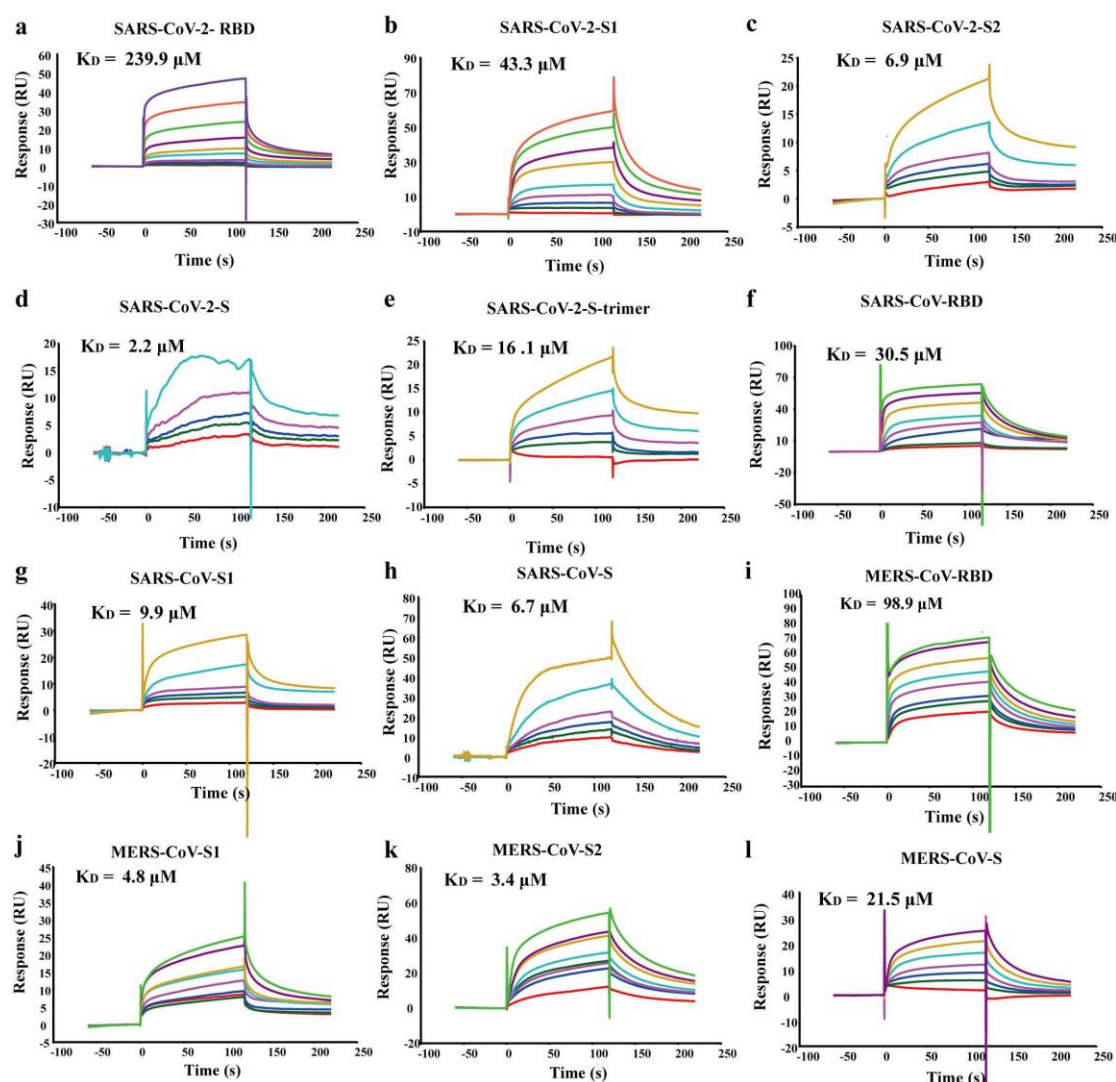
The oligosaccharides with relatively lower sulfation levels (**HS001-HS019**, the number of sulfate groups per monosaccharide unit  $<1.00$ ) have lower or almost no fluorescence signals. An analysis of the effect of the variation in sulfation revealed that the position of sulfation is another factor that strongly influence the binding. As shown in the Figure 4, removal of the 6-O-sulfate group from the glucosamine units significantly reduced the binding (compare **HS017** with **HS020**, **HS018** with **HS021**, and **HS019** with **HS022**), suggesting that the 6-O-sulfate in HS plays a crucial role in

determining its interaction with SARS-CoV-2-RBD. The importance of the 6-O-sulfate for binding is further supported by comparing the binding of **HS012**, **HS013**, **HS014**, **HS015**, and **HS016**, which shows that the one-by-one addition of sulfate to the 6-O-position of the glucosamine residues gradually increased the binding of HS with SARS-CoV-2-RBD. The microarray study also indicates that the binding is not positively related to the length of the HS chains. Short HS oligosaccharides could have comparable or even better binding properties (compare **HS001-HS006**, **HS007-HS0012**, and **HS020-HS024**).

The binding results of the SARS-CoV-2-S1 follow a similar trend as those of the SARS-CoV-2-RBD, with only minor differences for a few HS molecules like **HS001**, **HS007**, **HS016**, **HS021** and **HS024** (Fig. 4b). This is consistent with the assumption that RBD is the major determinant for viral S protein binding to HS.

Examination of the sequence of the SARS-CoV-2 S protein revealed the presence of a potential cleavage site for furin proteases (RRARS) at the S1/S2 boundary.<sup>9</sup> This is similar to the S protein of MERS-CoV, which also contain a furin consensus sequence RXXR (Fig. 2). Furin cleavage can occur in the secretory pathway of infected cells and breaks the covalent linkage between the S1 and S2 subunits.<sup>44</sup> The S2 subunit functions to fuse the virus to the host cell. It may interact with HS on the cell surface to facilitate the fusion process. Therefore, to determine the role of HS in SARS-CoV-2 infection, it is also important to investigate the binding of the S2 subunit to HS. Very interestingly, our results indicate that the SARS-CoV-2-S2 can also bind to HS, with the binding preferences similar to those of the SARS-CoV-2-RBD and the SARS-CoV-2-S1 (Fig. 4c). This suggests that HS may also play an important role during viral membrane fusion after the S1 subunit is removed from the S protein.

We also further investigated the binding full-length S protein and its trimer to HS using the same microarray. Our data showed that, although the binding to **HS020-HS024** still remains the highest, the full-length S protein has increased binding to HS with slightly less sulfation, particularly the molecules in groups **HS007-HS012** and **HS013-HS016** (Fig. 4d, 4e). Additionally, we found that the binding preferences of the RBDs, S1 subunits and full-length S proteins of SARS-CoV and MERS-CoV are similar to those of the subunits and S protein of SARS-CoV-2, although small differences are observed (Fig. 4f-k). Because the MERS-CoV S protein also contains a furin cleavage site, we studied the binding of its S2 subunit to the HS microarray. Once again, the results are similar to those of the S2 subunit of SARS-CoV-2. Overall, these data support the involvement of HS in the binding with the subunits and S proteins of SARS-CoV-2, SARS-CoV and MERS-CoV and support the importance of HS for the infection of these coronaviruses.

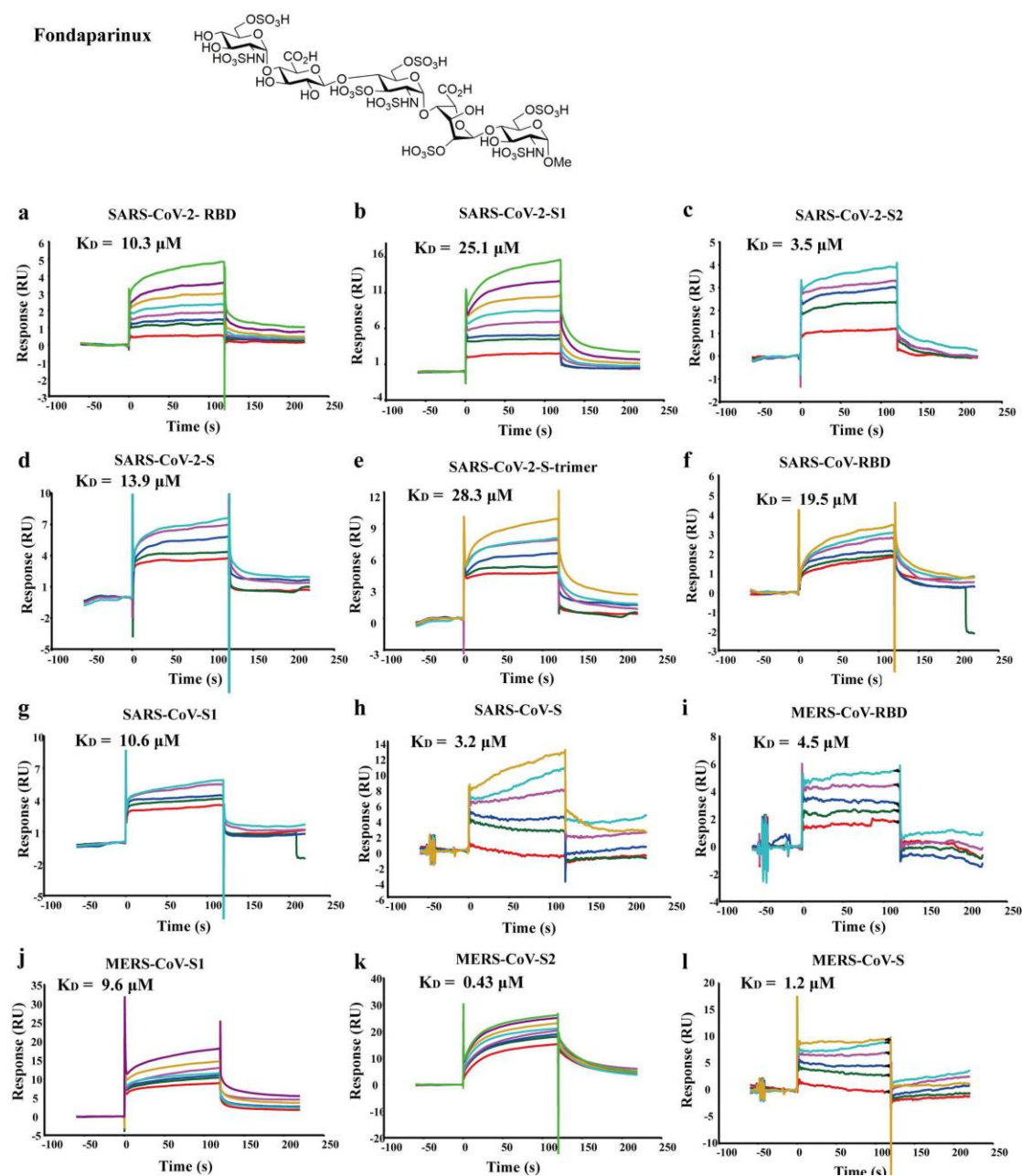


**Figure 5.** SPR sensorgrams showing the interactions between the long heparin molecule and immobilized proteins. The name and  $K_D$  value of each protein are presented on top of the corresponding sensorgrams.

**Measurement of the dissociation constants of a long heparin molecule and fondaparinux to the immobilized subunits and S proteins of SARS-CoV-2, SARS-CoV, and MERS-CoV by surface plasmon resonance (SPR).** To determine the relative binding strength of the subunits and S proteins of SARS-CoV-2, SARS-CoV, and MERS-CoV, we measured and compared the dissociation constants ( $K_D$ ) of the protein molecules studied in the microarray experiment using a real-time SPR-based binding assay. A commercially available porcine heparin from Sigma-Aldrich was used as the first binding partner for the measurement. It is a mixture of highly sulfated HS, with most chains in the molecular weight range of 17 to 19 kilodalton. This long heparin molecule is more similar to the HS molecules on the host cell surface than those on the microarray. In the SPR assay, protein molecules were covalently linked to the surface of the CM5 (carboxymethylated dextran) sensor chips by amine coupling. The heparin in various concentrations was then flowed over the immobilized proteins. The changes in refractive index by molecular interactions at the sensor surface were monitored and the dissociation constants were obtained by fitting the results using the



software available in the SPR instrument (Fig. 5).



**Figure 6.** The structure of fondaparinux and the SPR sensorgrams showing the interactions between fondaparinux and immobilized proteins. The name and  $K_D$  value of each protein are presented on top of the corresponding sensorgrams.

The comparison of the dissociation constants revealed that all the tested protein molecules can bind to the heparin, but their binding affinities are relatively low ( $K_D$ s are at the micromolar level). This agrees well with previous observations that HS is a weak binder to viral S proteins.<sup>45</sup> The results also showed that RBD has the lowest binding affinity among the tested protein molecules of SARS-CoV-2 (Fig. 5a). The S protein trimer and S1 subunit have relatively lower binding affinity in comparison with the full-length S protein and S2 subunit, respectively (Fig. 5b-e). Very similar trends were also

observed for differences in the binding affinities of the tested protein molecules of SARS-CoV and MERS-CoV (Fig. 5f-l).

In addition to the long heparin molecule, we also measured the binding of the protein molecules to fondaparinux (Arixtra®), which is an ultralow-molecular-weight-heparin (ULMWH) containing five monosaccharide units (Fig. 6). It has a well-defined chemical structure and is currently the only ULMWH that has been clinically approved as an anticoagulant. Fondaparinux is very similar to **HS023** in size, monosaccharide composition, and sulfation position and degree. The results showed that the  $K_D$  values of fondaparinux binding to the tested protein molecules are in the same range as those for the long heparin. This agrees well with the observation in the microarray study, suggesting that the length of the HS chains is not a determining factor for binding. Despite the similarity, there is a subtle but noticeable difference, which is that the binding affinities of the S1 subunit and RBD to fondaparinux are quite similar.

**Binding of the RBD, S1 subunit and S protein of SARS-CoV-2, and S proteins of SARS-CoV and MERS-CoV to sialylated glycan microarrays.** Many human viruses can interact with sialic acid-containing glycans present on the cell surface.<sup>17,20</sup> Such an interaction is normally mediated by the N-terminal domain of the S1 subunit.<sup>46</sup> In order to find out if SARS-CoV-2 can bind to sialic acid residues, we carried out microarray analyses of its S protein and subunits. The first microarray used contains 100 different N-glycans that may be found on the surface of cells. 49 of them are terminated with  $\alpha$ 2,3- and  $\alpha$ 2,6-linked sialic acid, also known as N-acetylneuraminic acid (Neu5Ac), 8 with  $\alpha$ 2,3- and  $\alpha$ 2,6-linked N-Glycolylneuraminic acid (Neu5Gc), and the rest with other glycan residues (Supporting Information, Table S1). The experiment was performed in the similar way as described for the HS microarray study. The microarray results showed that both the SARS-CoV-2-RBD and SARS-CoV-2-S1 gave no binding signal, suggesting that they may not be able to interact with sialylated N-glycans or the binding signal is too low to be detected. In order to confirm this finding, we also investigated the binding of the full-length S proteins of SARS-CoV-2, SARS-CoV and MERS-CoV to more sialylated glycans, including sialylated N- and O-linked glycans and glycolipid glycans (Supporting Information, Tables S2-4), but again no specific binding was detected.

## Discussion

For a virus like SARS-CoV-2 to establish infection, it must first attach itself to the surface of target cells in different organs and tissues. The S protein plays an essential role in this attachment process. Recently, the structure of SARS-CoV-2 S protein in the prefusion conformation was determined by the cryo-EM technique.<sup>10</sup> It shows that the overall structure of the SARS-CoV-2 S protein is very similar to that of the closely related SARS-CoV S protein, which is organized as a homotrimer. Each monomer can be divided into an N-terminal receptor-binding S1 subunit and a C-terminal fusion-mediating S2 subunit. The S1 subunits are located at the apex of the spike, making

them more accessible for binding to the proteinaceous receptor, ACE2. Although similar, there are some notable differences between the SARS-CoV-2 and SARS-CoV S proteins.<sup>8,10</sup> For example, the key amino acid residues involved in the binding of the SARS-CoV-2 S protein to ACE2 are largely different from those of SARS-CoV.<sup>9,47,48</sup> These differences may be related to the observed higher binding affinity of the SARS-CoV-2 S protein to ACE2.

Another important difference between the SARS-CoV-2 and SARS-CoV S proteins is that the former protein contains a multibasic protease recognition motif (RRARS) at the junction of S1 and S2.<sup>9</sup> A multibasic cleavage site (RSVRS) was also identified in the MERS-CoV S protein (Fig. 2).<sup>49</sup> The SARS-CoV S protein only has a monobasic amino acid (SSLRS) at the same site. The multibasic site can be processed by furin or related proprotein convertases, which are widely expressed in different tissues, before the virus is released from the host cell. By contrast, the monobasic site can be cleaved by TMPRSS2 or other cell-surface proteases (whose expression is confined to certain tissues) only after the virus is released from the host cell. It was reported that the cleavage at the junction of S1 and S2 activates the S protein for virus-cell fusion. Thus, the presence of the multibasic cleavage site may partially account for the enhanced infectivity and tropism of SARS-CoV-2 relative to SARS-CoV for human cells.<sup>10,50</sup> However, what is vague is how the virus attaches to the host cells after losing the S1 subunit, which is responsible for the binding to the proteinaceous receptor.

In addition to binding protein-based receptors, many viruses can interact with cell surface glycans, including GAGs and sialic acid-containing oligosaccharides. Depending on the virus, the glycan molecules can act as attachment factors, co-receptors or primary receptors.<sup>51</sup> Viruses typically bind GAGs through non-specific charge-based interactions. As one of the most abundant GAGs, HS appears to be the preferred binding partner for many viruses.<sup>33,52-55</sup> Sialic acids are normally terminal monosaccharide residues linked to glycans decorating cell surface glycoproteins, glycolipids, or other glycoconjugates.<sup>20,56</sup> In general, the interactions of viruses with HS or sialic acids are responsible for the first contact with host cells. Such contact may serve to concentrate viruses on the surface of target cells, facilitate their binding to more specific high-affinity protein receptors and/or promote their entry into host cells.<sup>57,58</sup> It has been demonstrated that virus binding and infection can be reduced by enzymatic removal HS or sialic acid from cell surface, or by treating virus with soluble HS or multivalent sialic acid conjugates.<sup>30,59-61</sup> Therefore, in order to better understand and treat COVID-19, it is necessary to carry out research to investigate the possible interactions between SARS-CoV-2 and HS and sialic acid-containing glycans in the forms of separate subunits and full-length proteins, and to assess if such interactions could represent a target for therapeutic intervention.

Similar to studies that have been successfully conducted for many other viruses, we used the microarray and SPR technology to study the binding of the RBDs, S1/S2

subunits and full-length S proteins of SARS-CoV-2, SARS-CoV and MERS-CoV, and a trimer of the SARS-CoV-2 S protein to HS and sialic acid.<sup>59</sup> The microarray results showed that all the tested protein molecules can bind to about half of the oligosaccharides on the HS microarray. In contrast, only background levels of fluorescence were detected on various sialylated glycan microarrays. This observation suggests all the tested protein molecules are able to bind to HS<sup>45,62</sup> and may not bind to and/or have very low binding affinity to sialylated glycans. This well agrees with previous studies showing that SARS-CoV can bind to HS<sup>62</sup> and the binding of the MERS-CoV S protein to sialic acid-containing glycans can only be detected after the S1 subunit was attached to a nanoparticle to enhance its avidity via multivalent interactions.<sup>17</sup>

Our results also suggested that the binding of the tested proteins to HS is related to HS sulfation position and degree. It seems that more 6-O-sulfate groups and higher sulfation degree normally lead to better binding. Because HSPGs exhibit different sulfation patterns in different tissues, such a binding specificity may contribute to the tropism of SARS-CoV-2 for human cells.<sup>63</sup> The length of HS appears not to be a critical factor for the binding. Short HS chains could have comparable binding specificity and signals. The comparison of the SPR  $K_D$  values obtained for the long heparin molecule and fondaparinux further support this finding. It implies that it may be possible to reduce the attachment of SARS-CoV-2 to the surface of host cells by low-molecular-weight-heparin (LMWH). This is in agreement with a recent study showing that LMWH treatment may be associated with better prognosis in some severe COVID-19 patients.<sup>64</sup> While these initial findings are encouraging, further research is required to determine if the binding to HS could affect the tropism and pathogenesis of SARS-CoV-2 and to determine if HS could be used for the inhibition of the infection of this virus.

Our SPR data of the tested protein molecules also showed that the S2 subunits could have similar or better binding affinity for HS as compared to those of the S1 subunits. This finding suggests that the cleaved S proteins of SARS-CoV-2 and MERS-CoV may depend on HS for interaction with the host cells during viral membrane fusion.

In parallel with our study, the Linhardt and Boons research teams also conducted studies to investigate the binding of S proteins to HS.<sup>45,65</sup> The absolute values of the dissociation constants determined in our experiments are largely different from those presented by these two teams. This can be seen from the binding of the full-length S protein of SARS-CoV-2, which was studied by all three teams (Supporting Information, Table S5). Our  $K_D$  is one order of magnitude higher than that reported by the Boons team, which in turn is three orders of magnitude higher than that reported by the Linhardt team. The differences in results may be due to the method of analysis and/or the experimental materials used in the studies. In our study, the tested proteins were immobilized on the surface of CM5 sensor chips, while in their experiments,

biotinylated heparin molecules were immobilized on the chips. At the same time, because the S glycoproteins and the heparin molecules were obtained from different sources, their composition may be different from each other.

In conclusion, through our study, we provided experimental evidence for whether or not the S protein of SARS-CoV-2 can bind to two types of cell-surface glycans, HS and sialic acid-containing glycans, which are commonly utilized by human viruses for attachment to target cells. Our data revealed that the SARS-CoV-2 S protein can weakly bind to HS in a sulfation-dependent manner. No binding with sialic acid residues was detected using the microarray assay. The results suggest that HS may act as an attachment factor to concentrate the virus at the cell surface and affect its tropism. Through comparison, we found that the S proteins of SARS-CoV and MERS-CoV have similar binding properties to HS as that of the SARS-CoV-2 S protein, indicating that HS binding may be a conserved feature for these three types of coronaviruses. Our data also revealed that the S2 subunits could bind equally well as the S1 subunits to HS. This binding may be an important element for viral attachment to the host cell surface after the removal of the N-terminal receptor-binding domains by protease cleavage. Overall, our findings support the potential importance of HS in SARS-CoV-2 infection and in the development of antiviral agents.

## Methods

**Reagents and cell lines.** High Five™ Cells for baculovirus expression were purchased from ThermoFisher Scientific and maintained in Express Five™ Medium. The recombinant SARS-CoV-2-S1 (16-685), SARS-CoV-2-S (16-1213, R683A, R685A), SARS-CoV-2-S-trimer (16-1213, R683A, R685A), SARS-CoV-RBD (306-527), SARS-CoV-S1 subunit (14-667), SARS-CoV-S (14-1195, R677A) with C-terminal His-tags were purchased from ACROBiosystems. The recombinant SARS-CoV-2-S2 (686-1213), MERS-CoV-RBD (367-606), MERS-CoV-S1 (1-725), MERS-CoV-S2 (726-1296), MERS-CoV-S (1-1297) with C-terminal His-tags were purchased from Sino Biological. The heparin sodium salt from porcine intestinal mucosa was purchased from Sigma-Aldrich/Merck. Fondaparinux (sodium salt) was purchased from Macklin Biochemical Co., Ltd (Shanghai, China). The glycan microarray experiments were performed by Creative Biochip Ltd (Nanjing, China).

**Expression and purification of SARS-CoV-2-RBD.** DNA containing the coding sequence for an N-terminal hemo signal peptide, the receptor binding domain (RBD, residues 319-541) of SARS-CoV-2 S protein and a C-terminal polyhistidine tag was amplified and inserted into a pFasebac1 vector for expression in High-5 insect cells using the Bac-to-Bac expression system (Invitrogen). The resulting recombinant protein, termed SARS-CoV-2-RBD-His, was secreted into cell culture medium, and subsequently purified on a nickel-nitrilotriacetic acid (Ni-NTA) affinity column, followed by a Superdex 200 gel filtration column (GE Healthcare). The final buffer for the protein



contains 10 mM Hepes (pH=7.0) and 100 mM NaCl. The purified SARS-CoV-2-RBD-His was concentrated to 3.5 mg/ml and flash frozen in liquid N<sub>2</sub> and stored at -80 degrees Celsius.

**Binding of recombinant proteins to glycan microarrays.** The tested protein molecules were first labeled with Cy3 fluorescent dye (10 mg/ml in DMSO). After dialysis, they were incubated at different concentrations with microarrays for 1 hour in the dark at room temperature. After incubation, the microarray slides were gently washed using washing buffer (20 mM Tris-Cl containing 0.1% tween 20, pH 7.4) to remove unbound proteins. Finally, the slides were scanned with a microarray scanner LuxScan-10K/Aat an excitation wavelength of 532 nm and evaluated by the Microarray Image Analyzer software.

**Measurement of dissociation constants by surface plasmon resonance (SPR).** The SPR measurements were performed using BIAcore T200 and S200 (GE healthcare). First, the carboxymethyl dextran matrix on CM5 sensor chip (GE Healthcare) was activated by injection of a 1:1 mixture of 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) and N-hydroxysuccinimide (NHS). Recombinant proteins in 10 mM acetate buffer (pH 4.5, GE Healthcare) was then injected over the chip surface at a flow rate of 30 µl/min to couple the amino groups of the recombinant proteins to the carboxymethyl dextran matrix. After the coupling reaction, the remaining activated ester groups were deactivated by ethanolamine. The binding study was carried out at 25°C in PBS-P buffer (GE Healthcare). The heparin molecule or fondaparinux at different concentrations were flowed over the immobilized recombinant proteins at a flow rate of 30 µl/min with a contact time of 60 s and a dissociation time of 100 s. The surface was regenerated by injection of 10 mM Glycine-HCl (pH 2.5) at a flow rate of 30 µl/min for 30 s. Data was collected and analyzed by BIA evaluation software (GE Healthcare).

## Author Contributions

All authors have given approval to the final version of the manuscript. ‡These authors contributed equally. W.H., B.M., S.C., Z.T. designed research; W.H., B.M., Z.L., X.W., X.G., Y.L., B.Q., S.S. performed research; W.H., B.M., S.C., Z.T. analyzed data; W.H., B.M., S.S., S.C., Z.T. wrote the paper.

## Acknowledgement

We would like to thank the National Natural Science Foundation of China (Grant number: 91853120), the National Major Scientific and Technological Special Project of China (Grant number: 2018ZX09711001-013), the National Key R&D Program of China (Grant number: 2018YFE0111400), the State Key Laboratory of Bioactive

Substance and Function of Natural Medicines, Institute of Materia Medica, and the Chinese Academy of Medical Sciences and Peking Union Medical College CRP-ICGEB Research Grant 2019 (Grant number: CRP/CHN19-02) for funding. We also thank Guangzhi Shan and Zhiling Zhu for their help in the SPR studies and Kai Zhang and Wenjie Zhou for their support to the research.

## Competing interests

The authors declare no competing interests.

## References

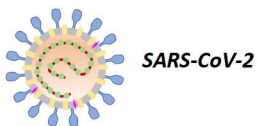
- 1 Zhu, N. *et al.* A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* **382**, 727-733 (2020).
- 2 Drosten, C. *et al.* Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N. Engl. J. Med.* **348**, 1967-1976 (2003).
- 3 Ksiazek, T. G. *et al.* A novel coronavirus associated with severe acute respiratory syndrome. *N. Engl. J. Med.* **348**, 1953-1966 (2003).
- 4 Zaki, A. M., van Boheemen, S., Bestebroer, T. M., Osterhaus, A. D. & Fouchier, R. A. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N. Engl. J. Med.* **367**, 1814-1820 (2012).
- 5 Chan, J. F. *et al.* A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet* **395**, 514-523 (2020).
- 6 Coronaviridae Study Group of the International Committee on Taxonomy of Viruses. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat. Microbiol.* **5**, 536-544 (2020).
- 7 Wang, H. *et al.* The genetic sequence, origin, and diagnosis of SARS-CoV-2. *Eur. J. Clin. Microbiol. Infect. Dis.* 1-7 (2020).
- 8 Lan, J. *et al.* Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **581**, 215-220 (2020).
- 9 Walls, A. C. *et al.* Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* **181**, 281-292.e286 (2020).
- 10 Wrapp, D. *et al.* Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**, 1260-1263 (2020).
- 11 Raj, V. S. *et al.* Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. *Nature* **495**, 251-254 (2013).
- 12 Ceccarelli, M., Berretta, M., Venanzi Rullo, E., Nunnari, G. & Cacopardo, B. Differences and similarities between Severe Acute Respiratory Syndrome (SARS)-CoronaVirus (CoV) and SARS-CoV-2. Would a rose by another name smell as sweet? *Eur. Rev. Med. Pharmacol. Sci.* **24**, 2781-2783 (2020).
- 13 SARS case fatality ratio, incubation period. Available from: [https://www.who.int/csr/sars/archive/2003\\_05\\_07a/en/](https://www.who.int/csr/sars/archive/2003_05_07a/en/) (2003).

- 14 Middle East respiratory syndrome coronavirus (MERS-CoV). Available from:  
[https://www.who.int/en/news-room/fact-sheets/detail/middle-east-respiratory-syndrome-coronavirus-\(mers-cov\)](https://www.who.int/en/news-room/fact-sheets/detail/middle-east-respiratory-syndrome-coronavirus-(mers-cov)) (2019).
- 15 Carette, J. E. *et al.* Ebola virus entry requires the cholesterol transporter Niemann-Pick C1. *Nature* **477**, 340-343 (2011).
- 16 Hallak, L. K., Kwilas, S. A. & Peeples, M. E. Cell surface  $\alpha$ 2,3-linked sialic acid facilitates Zika virus internalization. *Emerg. Microbes. Infect.* **8**, 426-437 (2019).
- 17 Li, W. *et al.* Identification of sialic acid-binding function for the Middle East respiratory syndrome coronavirus spike glycoprotein. *Proc. Natl. Acad. Sci. USA*. **114**, E8508-e8517, (2017).
- 18 Hallak, L. K. *et al.* Interaction between respiratory syncytial virus and glycosaminoglycans, including heparan sulfate. *Methods Mol. Biol.* **379**, 15-34 (2007).
- 19 Kalia, M., Chandra, V., Rahman, S. A., Sehgal, D. & Jameel, S. Interaction of Zika virus envelope protein with glycosaminoglycans. *Biochemistry* **56**, 1151-1162 (2017).
- 20 Trybala, E., Liljeqvist, J. A., Svennerholm, B. & Bergström, T. Structural basis for human coronavirus attachment to sialic acid receptors. *Nat. Struct. Mol. Biol.* **26**, 481-489 (2019).
- 21 Gomes, P. B. & Dietrich, C. P. Heparan sulfate proteoglycans are required for cellular binding of the hepatitis E virus ORF2 capsid protein and for viral infection. *J. Virol.* **83**, 12714-12724 (2009).
- 22 Toledo, O. M. & Dietrich, C. P. Herpes simplex virus types 1 and 2 differ in their interaction with heparan sulfate. *J. Virol.* **74**, 9106-9114 (2000).
- 23 Lyon, M. & Gallagher, J. T. Distribution of heparin and other sulfated glycosaminoglycans in vertebrates. *Comp. Biochem. Physiol. B.* **73**, 857-863 (1982).
- 24 Toledo, O. M. *et al.* Tissue specific distribution of sulfated mucopolysaccharides in mammals. *Biochim. Biophys. Acta.* **498**, 114-122 (1977).
- 25 Zautner, A. E., Jahn, B., Hammerschmidt, E., Wutzler, P. & Schmidtke, M. Bio-specific sequences and domains in heparan sulphate and the regulation of cell growth and adhesion. *Matrix. Biol.* **17**, 485-493 (1998).
- 26 Nader, H. B. *et al.* Heparan sulfates and heparins: similar compounds performing the same functions in vertebrates and invertebrates? *Braz. J. Med. Biol. Res.* **32**, 529-538 (1999).
- 27 Amonsén, M., Smith, D. F., Cummings, R. D. & Air, G. M. N- and 6-O-sulfated heparan sulfates mediate internalization of coxsackievirus B3 variant PD into CHO-K1 cells. *J. Virol.* **80**, 6629-6636 (2006).
- 28 Sasaki, M. *et al.* The role of heparan sulfate proteoglycans as an attachment factor for rabies virus entry and infection. *J. Infect. Dis.* **217**, 1740-1749 (2018).
- 29 Shukla, D. & Spear, P. G. Human parainfluenza viruses hPIV1 and hPIV3 bind oligosaccharides with  $\alpha$ 2-3-linked sialic acids that are distinct from those bound by H5 avian influenza virus hemagglutinin. *J. Virol.* **81**, 8341-8345 (2007).
- 30 Hilgard, P. & Stockert, R. Inhibition of viral adhesion and infection by sialic-acid-conjugated dendritic polymers. *Bioconjug. Chem.* **10**, 271-278 (1999).
- 31 Shukla, D. *et al.* Herpesviruses and heparan sulfate: an intimate relationship in aid of viral entry. *J. Clin. Invest.* **108**, 503-510 (2001).

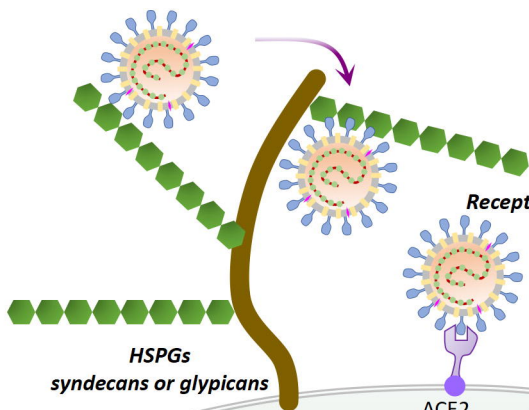
- 32 Hilgard, P. *et al.* Heparan sulfate proteoglycans initiate dengue virus infection of hepatocytes. *Hepatology* **32**, 1069-1077 (2000).
- 33 Tamhankar, M. *et al.* Heparan sulfate is an important mediator of Ebola virus infection in polarized epithelial cells. *J. Virol.* **15**, 135 (2018).
- 34 Guo, Y. R. *et al.* The origin, transmission and clinical therapies on coronavirus disease 2019 (COVID-19) outbreak - an update on the status. *Mil. Med. Res.* **7**, 11 (2020).
- 35 Wickramasinghe, I. N., de Vries, R. P., Gröne, A., de Haan, C. A. & Verheije, M. H. SARS-CoV-2 and SARS-CoV differ in their cell tropism and drug sensitivity profiles. doi:10.1101/2020.04.03.024257. *J. bioRxiv* (2020).
- 36 Hoffmann, M. *et al.* SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* **181**, 271-280.e278 (2020).
- 37 Stencel-Baerenwald, J. E., Reiss, K., Reiter, D. M., Stehle, T. & Dermody, T. S. The 2019 coronavirus (SARS-CoV-2) surface protein (spike) S1 receptor binding domain undergoes conformational change upon heparin binding. doi:10.1101/2020.02.29.971093. *J. bioRxiv* (2020).
- 38 Wickramasinghe, I. N. *et al.* Binding of avian coronavirus spike proteins to host factors reflects virus tropism and pathogenicity. *J. Virol.* **85**, 8903-8912, (2011).
- 39 Ahuja, S. Lectin microarray profiling and relative quantification of glycome associated with proteins of neonatal wt and rd1 mice retinæ. *Invest. Ophthalmol. Vis. Sci.* **54**, 3272-3280 (2013).
- 40 Stencel-Baerenwald, J. E. *et al.* The sweet spot: defining virus-sialic acid interactions. *Nat. Rev. Microbiol.* **12**, 739-749 (2014).
- 41 Corti, F. *et al.* N-terminal syndecan-2 domain selectively enhances 6-O heparan sulfate chains sulfation and promotes VEGFA(165)-dependent neovascularization. *Nat. Commun.* **10**, 1562 (2019).
- 42 Meneghetti, M. C. *et al.* Heparan sulfate and heparin interactions with proteins. *J. R. Soc. Interface.* **12**, 0589 (2015).
- 43 Zong, C. *et al.* Heparan sulfate microarray reveals that heparan sulfate-protein binding exhibits different ligand requirements. *J. Am. Chem. Soc.* **139**, 9534-9543 (2017).
- 44 Coutard, B. *et al.* The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade. *Antiviral. Res.* **176**, 104742 (2020).
- 45 Li, F., Li, W., Farzan, M. & Harrison, S. C. SARS-CoV-2 spike protein binds heparan sulfate in a length- and sequence-dependent manner. doi:10.1101/2020.05.10.087288. *J. bioRxiv* (2020).
- 46 Li, F. Receptor recognition mechanisms of coronaviruses: a decade of structural studies. *J. Virol.* **89**, 1954-1964 (2015).
- 47 Kleine-Weber, H., Elzayat, M. T., Hoffmann, M. & Pöhlmann, S. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. *Science* **309**, 1864-1868 (2005).
- 48 Ge, X. Y. *et al.* Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature* **503**, 535-538 (2013).
- 49 Raman, R., Tharakaraman, K., Sasisekharan, V. & Sasisekharan, R. Functional analysis of potential cleavage sites in the MERS-coronavirus spike protein. *Sci. Rep.* **8**, 16597 (2018).

- 50 Chang, A., Masante, C., Buchholz, U. J. & Dutch, R. E. Synthetic analogues of the snail toxin 6-bromo-2-mercaptotryptamine dimer (BrMT) reveal that lipid bilayer perturbation does not underlie its modulation of voltage-gated potassium channels. *Biochemistry* **57**, 2733-2743 (2018).
- 51 Raman, R. *et al.* Glycan-protein interactions in viral pathogenesis. *Curr. Opin. Struct. Biol.* **40**, 153-162 (2016).
- 52 Kesari, A. S., Sharkey, C. M. & Sanders, D. A. Human metapneumovirus (HMPV) binding and infection are mediated by interactions between the HMPV fusion protein and heparan sulfate. *J. Virol.* **86**, 3230-3243 (2012).
- 53 Mandl, C. W. *et al.* Adaptation of tick-borne encephalitis virus to BHK-21 cells results in the formation of multiple heparan sulfate binding sites in the envelope protein and attenuation in vivo. *J. Virol.* **75**, 5627-5637 (2001).
- 54 Kesari AS, Sharkey CM, Sanders DA. Role of heparan sulfate in entry and exit of Ross River virus glycoprotein-pseudotyped retroviral vectors. *Virology* **529**, 177-185 (2019).
- 55 Gao, H. *et al.* Role of heparan sulfate in the Zika virus entry, replication, and cell death. *Virology* **529**, 91-100 (2019).
- 56 Varki, A. *et al.* *Essentials of glycobiology*. (Cold Spring Harbor (NY): Cold spring harbor laboratory press; 2009. Chapter 14).
- 57 Aquino, R. S. & Park, P. W. Glycosaminoglycans and infection. *Front. Biosci.* **21**, 1260-1277 (2016).
- 58 Kamhi, E., Joo, E. J., Dordick, J. S. & Linhardt, R. J. Glycosaminoglycans in infectious disease. *Biol. Rev. Camb. Philos. Soc.* **88**, 928-943 (2013).
- 59 Chandra, N. *et al.* Sulfated glycosaminoglycans as viral decoy receptors for human adenovirus type 37. *Viruses* **11** (2019).
- 60 Techaarpornkul, S., Collins, P. L. & Peeples, M. E. Respiratory syncytial virus with the fusion protein as its only viral glycoprotein is less dependent on cellular glycosaminoglycans for attachment than complete virus. *Virology* **294**, 296-304 (2002).
- 61 Nilsson, E. C., Jamshidi, F., Johansson, S. M., Oberste, M. S. & Arnberg, N. Sialic acid is a cellular receptor for coxsackievirus A24 variant, an emerging virus with pandemic potential. *J. Virol.* **82**, 3061-3068 (2008).
- 62 Lang, J. *et al.* Inhibition of SARS pseudovirus cell entry by lactoferrin binding to heparan sulfate proteoglycans. *PLoS One* **6**, e23710 (2011).
- 63 Nagamine, S. *et al.* Organ-specific sulfation patterns of heparan sulfate generated by extracellular sulfatases Sulf1 and Sulf2 in mice. *J. Biol. Chem.* **287**, 9579-9590 (2012).
- 64 Tang, N. *et al.* Anticoagulant treatment is associated with decreased mortality in severe coronavirus disease 2019 patients with coagulopathy. *J. Thromb. Haemost.* **18**, 1094-1099 (2020).
- 65 Kim, S. Y. *et al.* Glycosaminoglycan binding motif at S1/S2 proteolytic cleavage site on spike glycoprotein may facilitate novel coronavirus (SARS-CoV-2) host cell entry. doi: <https://doi.org/10.1101/2020.04.14.041459>. *J. bioRxiv* (2020).





**Initial attachment**

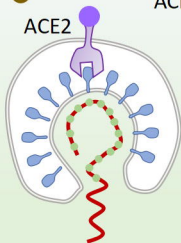


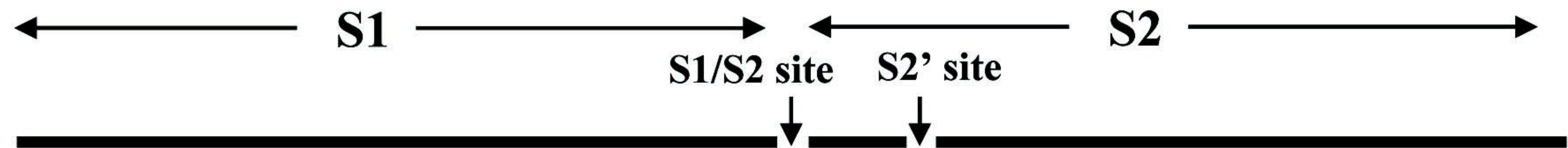
**ACE2**

**TMPRSS2**

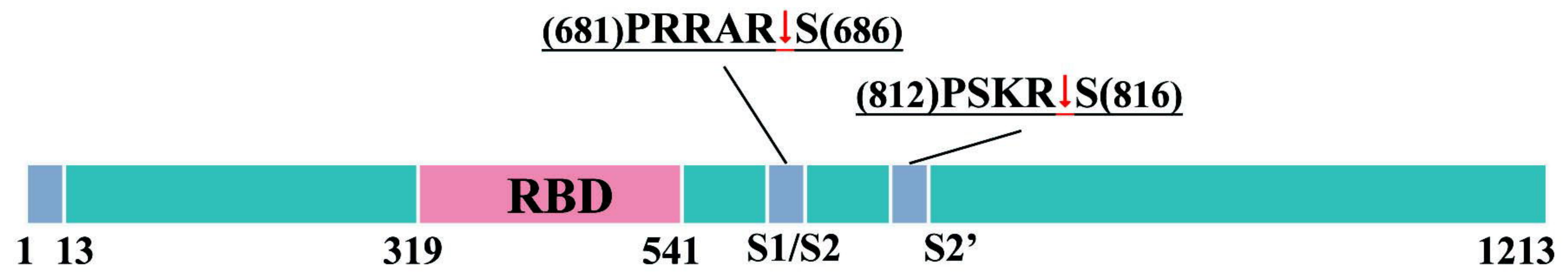
**ACE2**

**Entry & infection**

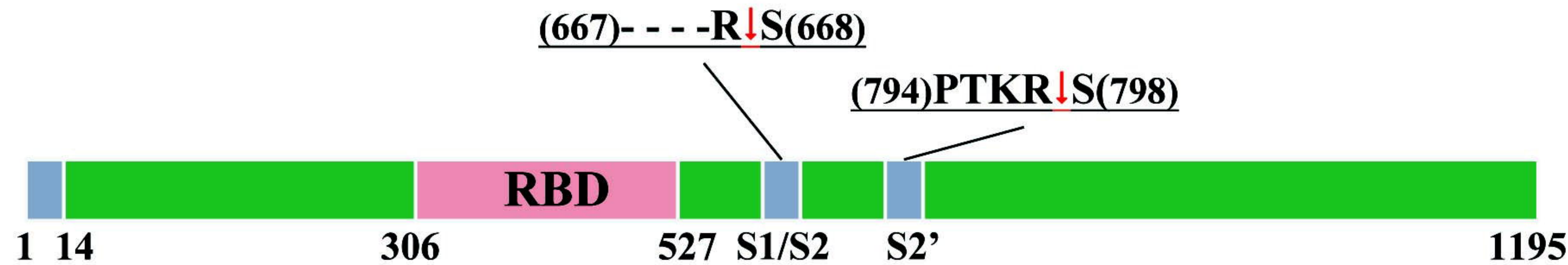




**SARS-CoV-2**



**SARS-CoV**



**MERS-CoV**

