Article

Binding of the SARS-CoV-2 spike protein to glycans

Wei Hao [a,1], Bo Ma [b,1], Ziheng Li [a], Xiaoyu Wang [b], Xiaopan Gao [a], Yaohao Li [b,c], Bo Qin [a], Shiying Shang [d], Sheng Cui [a,*], Zhongping Tan [b,*]

[a] NHC Key Laboratory of Systems Biology of Pathogens, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100730, China
[b] 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
[c] Department of Chemistry and Biochemistry and BioFrontiers Institute, University of Colorado, Boulder CO 80303, USA
[d] School of Pharmaceutical Sciences, Tsinghua University, Beijing 100084, China

ABSTRACT

The pandemic of the severe acute respiratory syndrome coronavirus 2 (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 glycolipids/glycoproteins. In this study, we examined and compared the binding of the subunits and spike (S) proteins of SARS-CoV-2, SARS-CoV, and Middle East respiratory disease (MERS)-CoV to these glycans. Our results revealed that the S proteins and subunits can bind to HS in a sulfation-dependent manner 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.

© 2021 Science China Press. Published by Elsevier B.V. and Science China Press. All rights reserved.

1. Introduction

The 2019 novel coronavirus (CoV) is the seventh coronavirus capable of infecting humans [1]. This vicious virus is changing 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 [2,3] and the Middle East respiratory disease (MERS) in 2012 [4]. In order to avoid a similar pandemic in future, it is necessary to develop a better understanding of these CoVs, especially with regard to effective ways to control the current COVID-19 pandemic and prevent the resurgences of the disease.

Studies have shown that the genome of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has about 80% nucleotide identity with that of SARS-CoV [5]. 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), whereas the nonstructural proteins (nsp1 to nsp16) are relatively more conserved. Based on genetic similarity, the 2019 novel CoV was named by the International Committee on Taxonomy of Viruses (ICTV) as SARS-CoV-2 [6]. SARS-CoV-2 is more genetically distant from MERS-CoV and shares only about 57% genome homology with MERS-CoV [5,7]. These similarities and differences are reflected in the interplay between viral surface glycoproteins and the cellular receptors. SARS-CoV-2 and SARS-CoV use the same receptor angiotensin-converting enzyme 2 (ACE2) to enter the target cells [8–10], whereas MERS-CoV uses dipeptidyl peptidase 4 (DPP4, also known as CD26) as the primary receptor [11]. However, there are many important findings that have not been explained by their genetic traits. For example, SARS-CoV-2 has a significantly lower mortality rate (2%) than SARS-CoV (15%) and MERS-CoV (35%), despite it is more transmissible among humans [12–14]. These findings suggest that more investigations, in addition to genetic analysis, are needed 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 [15,16]. Examples of such molecules include glycosaminoglycans (GAGs) and sialic acid-containing oligosaccharides [17–20]. GAGs are primarily localized at the outer surface of cells. Such a location makes them particularly suitable for acting as attachment factors to recruit

---

* Corresponding authors.
E-mail addresses: cui.sheng@ipb.pumc.edu.cn (S. Cui), zhongping.tan@imm.pumc.edu.cn (Z. Tan).
[1] These authors contributed equally to this work.

viruses to cell surfaces [21,22]. 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 [23,24]. The HS chains are mostly covalently linked as side chains to core proteins to form HS proteoglycans (HSPGs) (Fig. 1).

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 [25]. 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 [23,24,26]. 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 [27,28]. A similar trend was also observed for sialylation patterns of cell surface oligosaccharides [17,29,30]. For example, MERS-CoV S protein preferentially binds α2,3-linked sialic acids over α2,6-linked sialic acids on cell surface, and 5-N-glycolylation and 7,9-O-acetylation of sialic acids disrupts their interaction. These findings implicate that a possible relationship may exist between the distribution of different types of HS/sialylated glycans and the viral tropism [31–33]. 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 completely correlated with the ACE2 expression [34,35]. Some other factors, such as proteases and glycans, may be the determinant of cellular susceptibility to the infection with this virus [36]. 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 [37]. 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 [38].

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 might not be a very critical factor for the binding. 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.

2. Materials and methods

2.1. Reagents and cell lines

High Five™ cells for baculovirus expression were purchased from ThermoFisher Scientific (Shanghai, China) and maintained in Express Five™ Medium. The recombinant SARS-CoV-2-S1
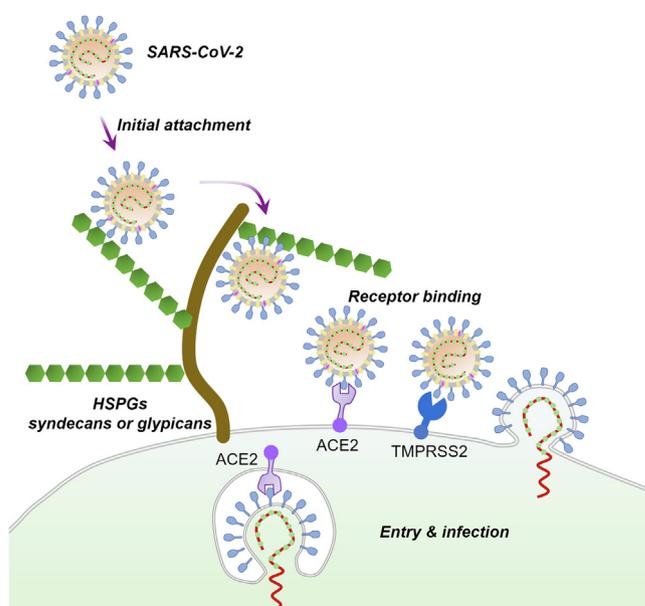

Fig. 1. A possible mechanism for SARS-CoV-2 entry and infection. At the early stage of the infection process, SARS-CoV-2 may first interact with the HSPGs on the surface of susceptible cells using the S protein protruding from the virus particle. This initial attachment may promote the subsequent binding of the virus to the high-affinity entry receptor ACE2. The transmembrane protease serine 2 (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.

(16–685), SARS-CoV-2-S (16–1213, R683A, and R685A), SARS-CoV-2-S-trimer (16–1213, R683A, and R685A), SARS-CoV-RBD (306–527), SARS-CoV-S1 subunit (14–667), SARS-CoV-S (14–1195 and R677A) with C-terminal His-tags were purchased from ACROBiosystems (Beijing, China). The recombinant SARS-CoV-2-S2 (686–1213), MERS-CoV-RBD (367–606), MERS-CoV-S1 (1–725), MERS-CoV-S2 (726–1296), and MERS-CoV-S (1-1297) with C-terminal His-tags were purchased from Sino Biological (Beijing, China). The heparin sodium salt from porcine intestinal mucosa was purchased from Sigma-Aldrich/Merck (Shanghai, China). Fondaparinux (sodium salt, also called Arixtra®) was purchased from Macklin Biochemical Co., Ltd (Shanghai, China). The glycan microarray experiments were performed by Creative Biochip Ltd (Nanjing, China).

2.2. 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 Five™ cells using the Bac-to-Bac expression system (ThermoFisher Scientific). 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, Shanghai, China). The final buffer for the protein contains 10 mmol/L Hepes (pH 7.0) and 100 mmol/L NaCl. The purified SARS-CoV-2-RBD-His was concentrated to 3.5 mg/mL and flash frozen in liquid N_2 and stored at −80°C.

2.3. Binding of recombinant proteins to glycan microarrays

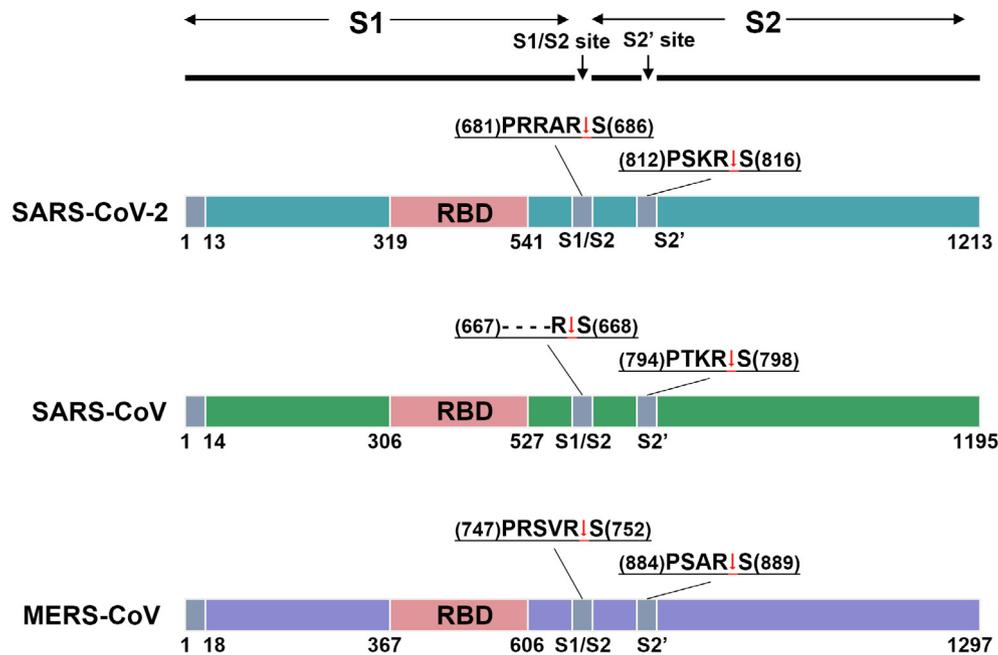The microarray experiments were carried out using the protocol described by the National Center for Functional Glycomics (https://

Fig. 2. Schematic representation of CoV S proteins. The diagrams show the domain organization of the S proteins of SARS-CoV-2, SARS-CoV, and MERS-CoV. The sequences of the S1/S2 and S2' cleavage sites are given on top of each diagram. The arrows indicate the cleave points.

ncfg.hms.harvard.edu/protocols). Briefly, the tested protein molecules were first labeled with Cy3 fluorescent dye (10 mg/mL in dimethyl sulfoxide (DMSO)). After dialysis, they were incubated at different concentrations with microarrays for 1 h in the dark at room temperature. After incubation, the microarray slides were gently washed using washing buffer (20 mmol/L Tris-HCl, 150 mmol/L NaCl, 2 mmol/L CaCl$_2$, 2 mmol/L MgCl$_2$, and 0.05% Tween-20, pH 7.4) to remove unbound proteins. Finally, the slides were scanned with a microarray scanner LuxScan-10 K/A at an excitation wavelength of 532 nm and evaluated by the Microarray Image Analyzer software (Version CB-001, Creative Biochip Ltd).

2.4. Measurement of dissociation constants by surface plasmon resonance (SPR)

The SPR measurements were performed using BIAcore S200 (Cytiva, Marlborough, USA) following a protocol provided by the manufacturer. First, the carboxymethyl dextran matrix on CM5 sensor chip (Cytiva) 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 mmol/ L 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℃ in phosphate buffer saline (PBS, 20 mmol/L phosphate buffer with 2.7 mmol/L KCl, 137 mmol/L NaCl, and 0.05% Tween-20, pH 7.4). The heparin molecule or fonda-parinux at different concentrations were flowed over the immobi-lized 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 mmol/L glycine-HCl (pH 2.5) at a flow rate of 30 μL/min for 30 s. Data was collected and ana-lyzed using the BIA evaluation software (Version 1.0, Cytiva).

3. Results

3.1. 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 pro-tein might bind heparin [9]. In order to determine if there is any preference of the SARS-CoV-2 S protein for particular HS struc-tures, 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 sys-tematic differences in their length, monosaccharide composition, and sulfation pattern (Fig. 3). The microarray experiment was per-formed using a previously established standard protocol [29,39,40]. 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 μg/mL. Increas-ing the concentration of the protein was not found to produce noticeable changes in binding (Fig. S1 online).

Quantification of fluorescence revealed that the SARS-CoV-2-RBD is able to bind to almost half of the molecules on the microar-ray, 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 [41–43]. As shown in Fig. 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 monosac-charide), which is followed by those of HS021 and HS024 (1.25 sul-fate 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 com-position (compare HS020 with HS022, and HS021 with HS024).
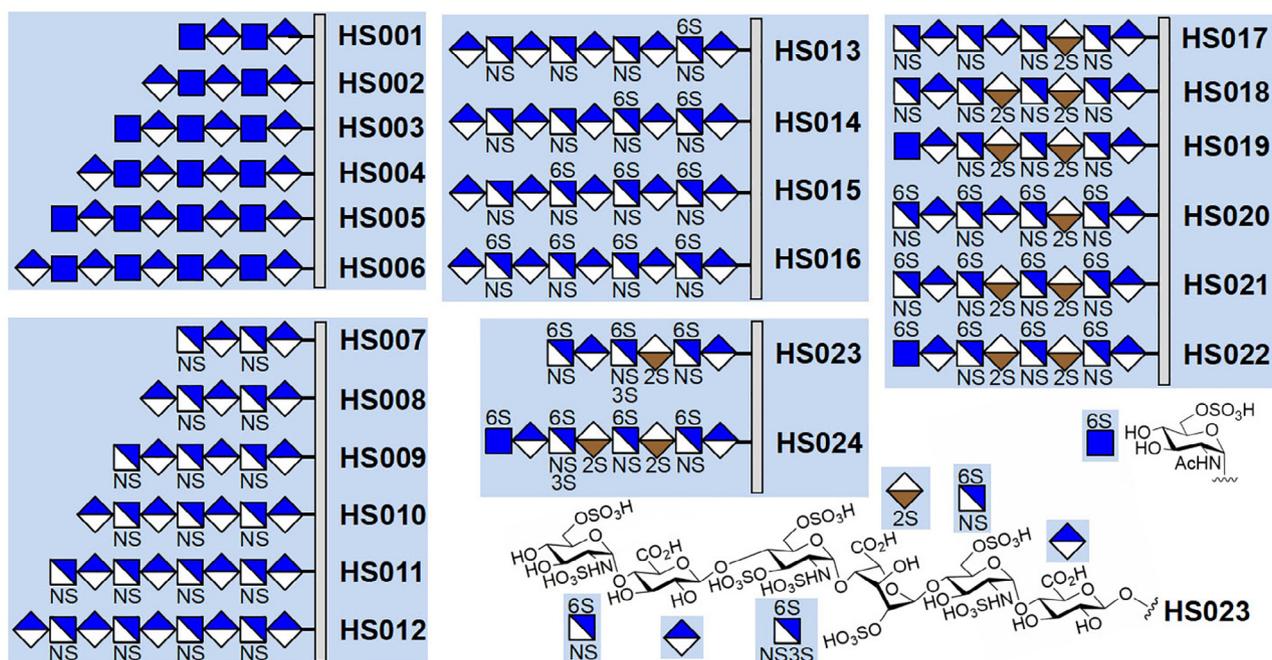
Fig. 3. The numbering and structures of the HS oligosaccharides on the microarray. Each HS chain is covalently attached to the microarray slide via the reducing end. The HS oligosaccharides can be roughly divided into five groups according to their differences in the monosaccharide compositions and sulfation patterns.

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 Fig. 4, the one-by-one addition of sulfate to the 6-O-position of the glucosamine residues gradually increased the binding (compare HS012, HS013, HS014, HS015, and HS016), suggesting that the 6-O-sulfate plays an important role in determining the interaction between HS and SARS-CoV-2-RBD. The importance of the 6-O-sulfation is further supported by comparing the binding of HS017–HS019 with HS020–HS022, which shows that addition of 4 6-O-sulfate groups to HS017–HS019 significantly increase the binding of HS. Moreover, the results also suggest that SARS-CoV-2-RBD may prefer to bind certain 6-O-sulfated sequences, such as IdoA2S-GlcNS6S (compare HS016 and HS021).

The microarray study of these 24 oligosaccharides indicates that the binding might not be positively related to the length of the HS chains. Shorter HS molecules 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 [9]. 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 [44]. 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, e). 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.

3.2. 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 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 kD. 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
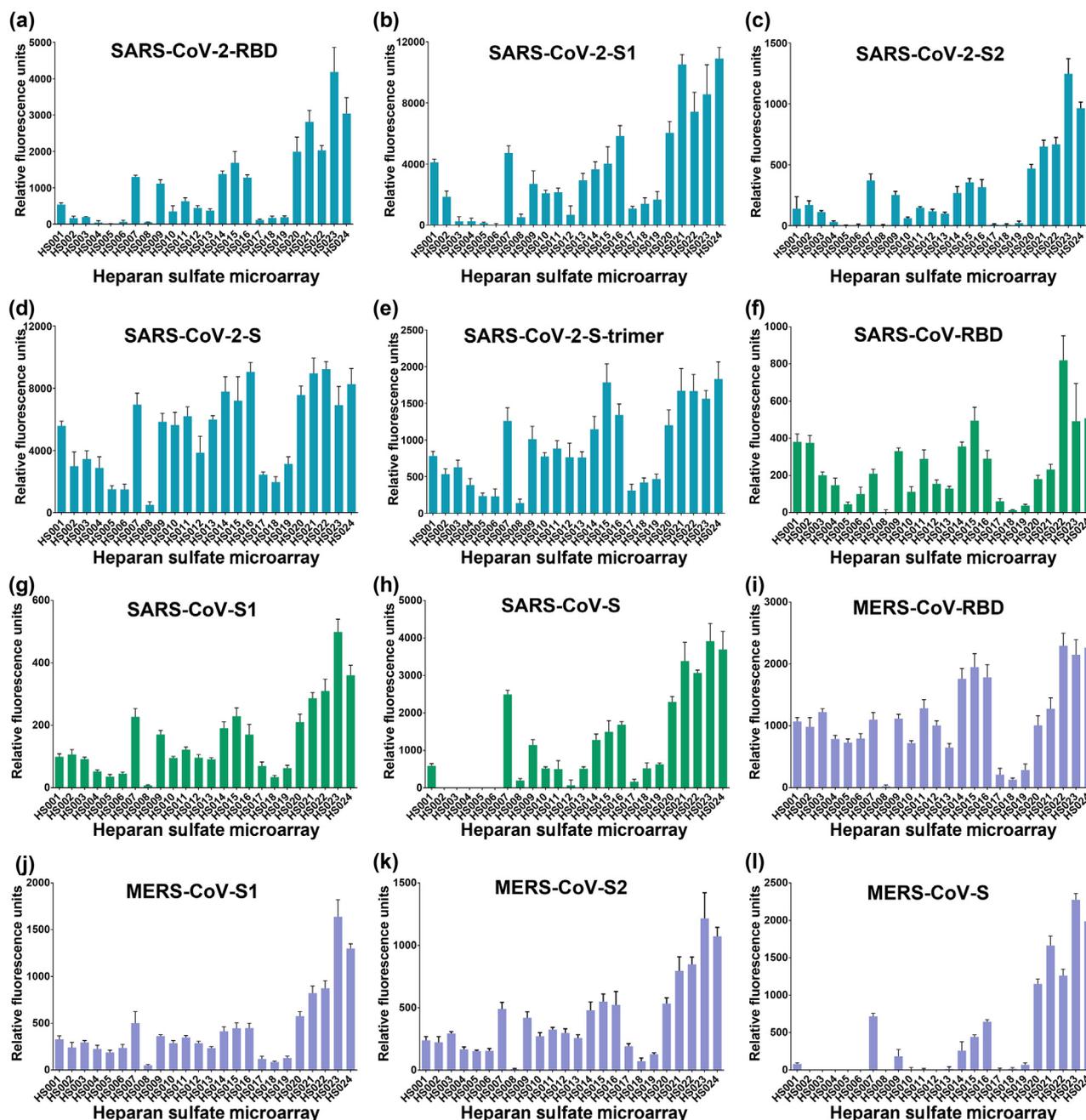
Fig. 4. The binding of full-length CoV S proteins and their domains to HS microarrays. (a) SARS-CoV-2-RBD binding at the concentration of 2.5 μg/mL, (b) SARS-CoV-2-S1 at the concentration of 2 μg/mL, (c) SARS-CoV-2-S2 at the concentration of 2 μg/mL, (d) SARS-CoV-2-S at the concentration of 2 μg/mL, (e) SARS-CoV-2-trimer at the concentration of 2 μg/mL, (f) SARS-CoV-RBD at the concentration of 2 μg/mL, (g) SARS-CoV-S1 at the concentration of 2 μg/mL, (h) SARS-CoV-2-S at the concentration of 2 μg/mL, (i) MERS-CoV-RBD at the concentration of 4 μg/mL, (j) MERS-CoV-S1 at the concentration of 4 μg/mL, (k) MERS-CoV-S2 at the concentration of 2 μg/mL, (l) MERS-CoV-S at the concentration of 2 μg/mL. The fluorescence intensity was measured at an excitation wavelength of 532 nm. All error bars are standard deviation of more than three replicates.

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).

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 [45]. 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, 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
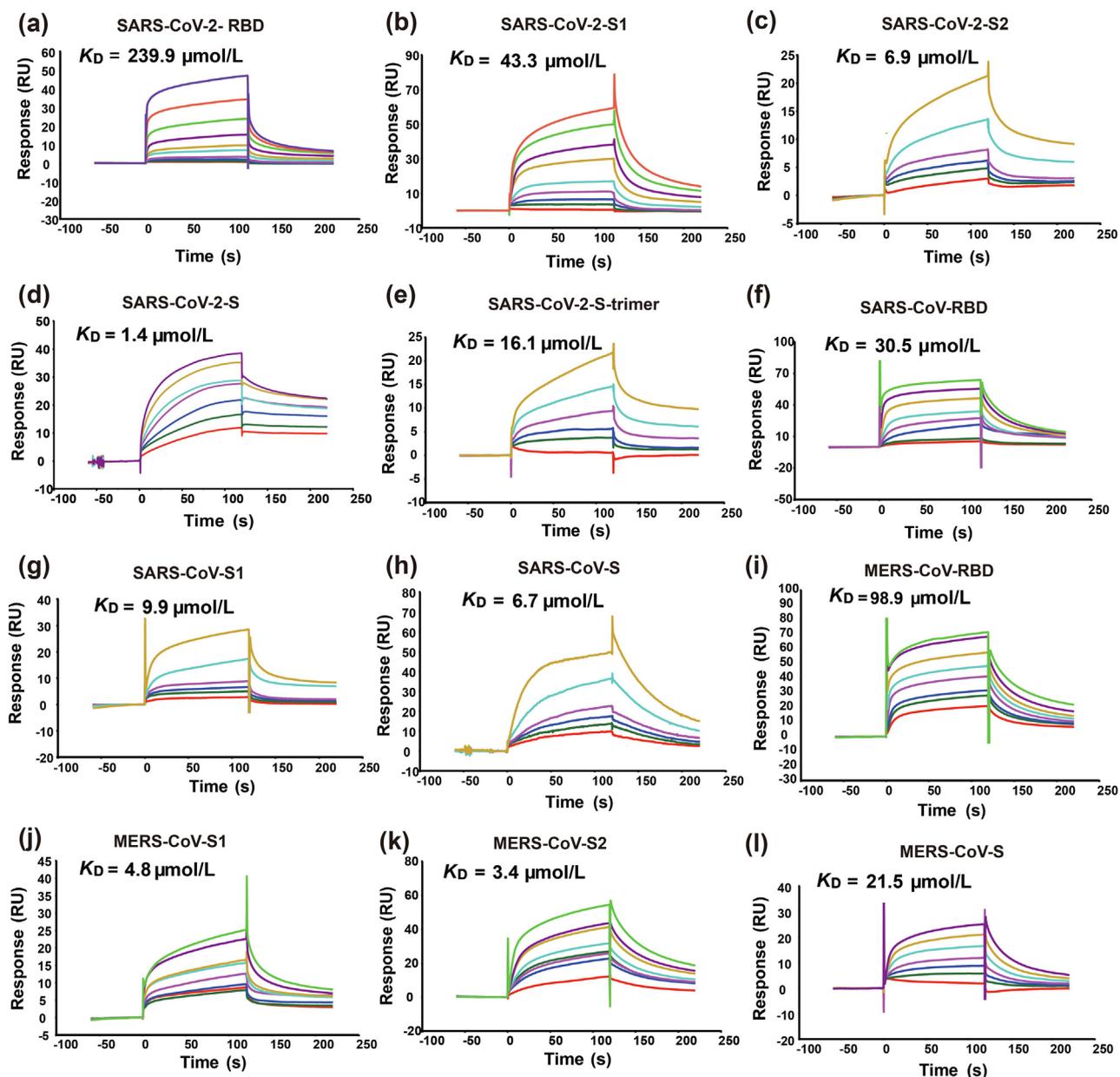
Fig. 5. SPR sensorgrams of the interactions between a commercially available porcine heparin and immobilized S proteins and their domains. Panels (a−l) show the binding of heparin to SARS-CoV-2-RBD, S1, S2, S, trimer, SARS-CoV-RBD, S1, S, and MERS-CoV-RBD, S1, S2, S, respectively. The name of each protein and its K_D value are presented on top of the corresponding sensorgram. RU, resonance unit.

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.

3.3. 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 [17,20]. Such an interaction is normally mediated by the N-terminal domain of the S1 subunit [46]. 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. Of these, 49 are terminated with α2,3- and α2,6-linked sialic acid, also known as N-acetylneuraminic acid (Neu5Ac), 8 with α2,3- and α2,6-linked N-Glycolylneuraminic acid (Neu5Gc), and the rest with other glycan residues (Table S1 online). 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 (Tables S2–S4 online), but again no specific binding was detected.
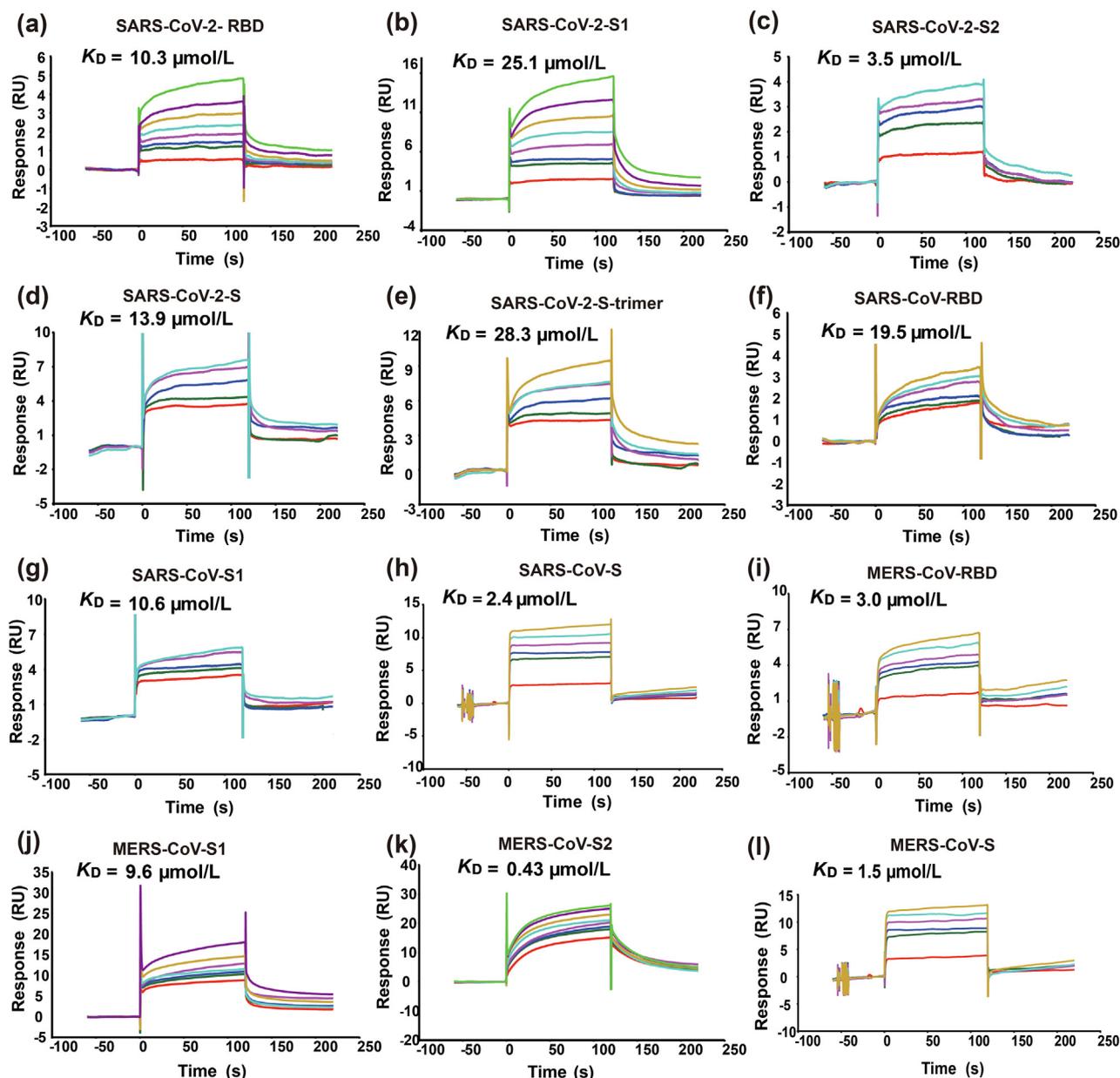
Fig. 6. SPR sensorgrams of the interactions between fondaparinux and immobilized S proteins and their domains. Panels (a–l) are the binding of fondaparinux to SARS-CoV-2-RBD, S1, S2, S, trimer, SARS-CoV-RBD, S1, S, and MERS-CoV-RBD, S1, S2, S, respectively. The name and K_D value of each protein are presented on top of the corresponding sensorgram. RU, resonance unit.

4. Discussion and conclusion

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-electron microscopy (cryo-EM) technique [10]. 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 [8,10].

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 [9,47,48]. 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 [9]. A multibasic cleavage site (RSVRS) was also identified in the MERS-CoV S protein (Fig. 2) [49]. 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 transmembrane protease serine 2 (TMPRSS2) or other cell-surface proteases (whose expression is confined to certain tis-

sues) 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 [10,50]. 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 [51]. Viruses typically bind GAGs through nonspecific charge-based interactions. As one of the most abundant GAGs, HS appears to be the preferred binding partner for many viruses [33,52–55]. Sialic acids are normally terminal monosaccharide residues linked to glycans decorating cell surface glycoproteins, glycolipids, or other glycoconjugates [20,56]. 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 [57,58]. 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 [30,59–61]. 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 [59]. 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 [45,62] 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 [62] 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 [17].

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 [63]. 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 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 [64]. 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 [45,65]. 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 (Table S5 online). 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. We carried out further research to identify possible causes for the differences. Our results suggested that the discrepancy between our K_D values and those reported by the Linhardt and Boons teams may be due to the method of analysis and/or the experimental materials used in the studies (Figs. S2, S3, and Table S6 online). 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, thus leading to further differences.

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.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgments

(Institute of Medicinal Biotechnology, Chinese Academy of Medical Sciences) for their help in the SPR studies, and Kai Zhang (Portever Shipping Co., Ltd.) and Wenjie Zhou (Tianjin Tongquan International Logistics Co., Ltd.) for their support to the research.

Author contributions

Appendix A. Supplementary materials

Supplementary materials to this article can be found online at https://doi.org/10.1016/j.scib.2021.01.010.

References

[1] Zhu N, Zhang D, Wang W, et al. A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med 2020;382:727–33.

[2] Drosten C, Günther S, Preiser W, et al. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. N Engl J Med 2003;348:1967–76.

[3] Ksiazek TG, Erdman D, Goldsmith CS, et al. A novel coronavirus associated with severe acute respiratory syndrome. N Engl J Med 2003;348:1953–66.

[4] Zaki AM, van Boheemen S, Bestebroer TM, et al. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. N Engl J Med 2012;367:1814–20.

[5] Chan JF, Yuan S, Kok KH, 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 2020;395:514–23.

[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 2020; 5:536–44.

[7] Wang H, Li X, Li T, et al. The genetic sequence, origin, and diagnosis of SARS-CoV-2. Eur J Clin Microbiol Infect Dis 2020;39:1629–35.

[8] Lan J, Ge J, Yu J, et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. Nature 2020;581:215–20.

[9] Walls AC, Park YJ, Tortorici MA, et al. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. Cell 2020;181:281–292.e6.

[10] Wrapp D, Wang N, Corbett KS, et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science 2020;367:1260–3.

[11] Raj VS, Mou H, Smits SL, et al. Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. Nature 2013;495:251–4.

[12] Ceccarelli M, Berretta M, Venanzi Rullo E, et al. 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 2020;24:2781–3.

[13] SARS case fatality ratio, incubation period. https://www.who.int/csr/sars/archive/2003_05_07a/en/ [Accessed May 2003].

[14] Middle East respiratory syndrome coronavi rus (MERS-CoV). https://www.who.int/en/news-room/fact-sheets/detail/middle-east-respiratory-syndrome-coronavirus-(mers-cov) [Accessed March 2009].

[15] Carette JE, Raaben M, Wong AC, et al. Ebola virus entry requires the cholesterol transporter Niemann-Pick C1. Nature 2011;477:340–3.

[16] Tan CW, Huan Hor CH, Kwek SS, et al. Cell surface α2,3-linked sialic acid facilitates Zika virus internalization. Emerg Microbes Infec 2019;8:426–37.

[17] Li W, Hulswit RJG, Widjaja I, et al. Identification of sialic acid-binding function for the Middle East respiratory syndrome coronavirus spike glycoprotein. Proc Natl Acad Sci USA 2017;114:E8508–17.

[18] Hallak LK, Kwilas SA, Peeples ME. Interaction between respiratory syncytial virus and glycosaminoglycans, including heparan sulfate. Methods Mol Biol 2007;379:15–34.

[19] Kim SY, Zhao J, Liu X, et al. Interaction of Zika virus envelope protein with glycosaminoglycans. Biochemistry 2017;56:1151–62.

[20] Tortorici MA, Walls AC, Lang Y, et al. Structural basis for human coronavirus attachment to sialic acid receptors. Nat Struct Mol Biol 2019;26:481–9.

[21] Kalia M, Chandra V, Rahman SA, et al. Heparan sulfate proteoglycans are required for cellular binding of the hepatitis E virus ORF2 capsid protein and for viral infection. J Virol 2009;83:12714–24.

[22] Trybala E, Liljeqvist JA, Svennerholm B, et al. Herpes simplex virus types 1 and 2 differ in their interaction with heparan sulfate. J Virol 2000;74:9106–14.

[23] Gomes PB, Dietrich CP. Distribution of heparin and other sulfated glycosaminoglycans in vertebrates. Comp Biochem Physiol B 1982;73:857–63.

[24] Toledo OM, Dietrich CP. Tissue specific distribution of sulfated mucopolysaccharides in mammals. Biochim Biophys Acta 1977;498:114–22.

[25] Lyon M, Gallagher JT. Bio-specific sequences and domains in heparan sulphate and the regulation of cell growth and adhesion. Matrix Biol 1998;17:485–93.

[26] Nader HB, Chavante SF, dos-Santos EA, et al. Heparan sulfates and heparins: similar compounds performing the same functions in vertebrates and invertebrates?. Braz J Med Biol Res 1999;32:529–38.

[27] Zautner AE, Jahn B, Hammerschmidt E, et al. N- and 6-O-sulfated heparan sulfates mediate internalization of coxsackievirus B3 variant PD into CHO-K1 cells. J Virol 2006;80:6629–36.

[28] Sasaki M, Anindita PD, Ito N, et al. The role of heparan sulfate proteoglycans as an attachment factor for rabies virus entry and infection. J Infect Dis 2018;217:1740–9.

[29] Amonsen M, Smith DF, Cummings RD, et al. Human parainfluenza viruses hPIV1 and hPIV3 bind oligosaccharides with α2-3-linked sialic acids that are distinct from those bound by H5 avian influenza virus hemagglutinin. J Virol 2007;81:8341–5.

[30] Reuter JD, Myc A, Hayes MM, et al. Inhibition of viral adhesion and infection by sialic-acid-conjugated dendritic polymers. Bioconjug Chem 1999;10:271–8.

[31] Shukla D, Spear PG. Herpesviruses and heparan sulfate: an intimate relationship in aid of viral entry. J Clin Invest 2001;108:503–10.

[32] Hilgard P, Stockert R. Heparan sulfate proteoglycans initiate dengue virus infection of hepatocytes. Hepatology 2000;32:1069–77.

[33] Tamhankar M, Gerhardt DM, Bennett RS, et al. Heparan sulfate is an important mediator of Ebola virus infection in polarized epithelial cells. Virol J 2018;15:135.

[34] Guo YR, Cao QD, Hong ZS, et al. The origin, transmission and clinical therapies on coronavirus disease 2019 (COVID-19) outbreak - an update on the status. Mil Med Res 2020;7:11.

[35] Bojkova D, McGreig JE, McLaughlin K-M, et al. SARS-CoV-2 and SARS-CoV differ in their cell tropism and drug sensitivity profiles. bioRxiv 2020. https://doi.org/10.1101/2020.04.03.024257.

[36] Hoffmann M, Kleine-Weber H, Schroeder S, et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. Cell 2020;181:271–80. e278.

[37] Mycroft-West C, Su D, Elli S, et al. The 2019 coronavirus (SARS-CoV-2) surface protein (spike) S1 receptor binding domain undergoes conformational change upon heparin binding. bioRxiv 2020. https://doi.org/10.1101/2020.02.29.971093.

[38] Wickramasinghe IN, de Vries RP, Gröne A, et al. Binding of avian coronavirus spike proteins to host factors reflects virus tropism and pathogenicity. J Virol 2011;85:8903–12.

[39] Ahuja S. Lectin microarray profiling and relative quantification of glycome associated with proteins of neonatal *wt* and *rd1* mice retinae. Invest Ophthalmol Vis Sci 2013;54:3272–80.

[40] Stencel-Baerenwald JE, Reiss K, Reiter DM, et al. The sweet spot: defining virus-sialic acid interactions. Nat Rev Microbiol 2014;12:739–49.

[41] Corti F, Wang Y, Rhodes JM, et al. N-terminal syndecan-2 domain selectively enhances 6-O heparan sulfate chains sulfation and promotes VEGFA$_{165}$-dependent neovascularization. Nat Commun 2019;10:1562.

[42] Meneghetti MC, Hughes AJ, Rudd TR, et al. Heparan sulfate and heparin interactions with proteins. J R Soc Interface 2015;12:0589.

[43] Zong C, Venot A, Li X, et al. Heparan sulfate microarray reveals that heparan sulfate-protein binding exhibits different ligand requirements. J Am Chem Soc 2017;139:9534–43.

[44] Coutard B, Valle C, de Lamballerie X, 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 2020;176:104742.

[45] Liu L, Chopra P, Li X, et al. SARS-CoV-2 spike protein binds heparan sulfate in a length- and sequence-dependent manner. bioRxiv 2020. https://doi.org/10.1101/2020.05.10.087288.

[46] Li F. Receptor recognition mechanisms of coronaviruses: a decade of structural studies. J Virol 2015;89:1954–64.

[47] Li F, Li W, Farzan M, et al. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. Science 2005;309:1864–8.

[48] Ge XY, Li JL, Yang XL, et al. Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. Nature 2013;503:535–8.

[49] Kleine-Weber H, Elzayat MT, Hoffmann M, et al. Functional analysis of potential cleavage sites in the MERS-coronavirus spike protein. Sci Rep 2018;8:16597.

[50] Dockendorff C, Gandhi DM, Kimball IH, et al. 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 2018;57:2733–43.

[51] Raman R, Tharakaraman K, Sasisekharan V, et al. Glycan-protein interactions in viral pathogenesis. Curr Opin Struct Biol 2016;40:153–62.

[52] Chang A, Masante C, Buchholz UJ, et al. Human metapneumovirus (HMPV) binding and infection are mediated by interactions between the HMPV fusion protein and heparan sulfate. J Virol 2012;86:3230–43.

[53] Mandl CW, Kroschewski H, Allison SL, 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 2001;75:5627–37.

[54] Kesari AS, Sharkey CM, Sanders DA. Role of heparan sulfate in entry and exit of Ross River virus glycoprotein-pseudotyped retroviral vectors. Virology 2019;529:177–85.

[55] Gao H, Lin Y, He J, et al. Role of heparan sulfate in the Zika virus entry, replication, and cell death. Virology 2019;529:91–100.

[56] Varki A, Cummings R D, Esko J D, et al. Essentials of glycobiology. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2009.

[57] Aquino RS, Park PW. Glycosaminoglycans and infection. Front Biosci (Landmark Ed) 2016;21:1260–77.

[58] Kamhi E, Joo EJ, Dordick JS, et al. Glycosaminoglycans in infectious disease. Biol Rev Camb Philos Soc 2013;88:928–43.

[59] Chandra N, Liu Y, Liu JX, et al. Sulfated glycosaminoglycans as viral decoy receptors for human adenovirus type 37. Viruses 2019;11:247.

[60] Techaarpornkul S, Collins PL, Peeples ME. Respiratory syncytial virus with the fusion protein as its only viral glycoprotein is less dependent on cellular glycosaminoglycans for attachment than complete virus. Virology 2002;294:296–304.

[61] Nilsson EC, Jamshidi F, Johansson SM, et al. Sialic acid is a cellular receptor for coxsackievirus A24 variant, an emerging virus with pandemic potential. J Virol 2008;82:3061–8.

[62] Lang J, Yang N, Deng J, et al. Inhibition of SARS pseudovirus cell entry by lactoferrin binding to heparan sulfate proteoglycans. PLoS One 2011;6:e23710.

[63] Nagamine S, Tamba M, Ishimine H, et al. Organ-specific sulfation patterns of heparan sulfate generated by extracellular sulfatases Sulf1 and Sulf2 in mice. J Biol Chem 2012;287:9579–90.

[64] Tang N, Bai H, Chen X, et al. Anticoagulant treatment is associated with decreased mortality in severe coronavirus disease 2019 patients with coagulopathy. J Thromb Haemost 2020;18:1094–9.

[65] Kim SY, Jin W, Sood A, et al. Glycosaminoglycan binding motif at s1/s2 proteolytic cleavage site on spike glycoprotein may facilitate novel coronavirus (SARS-CoV-2) host cell entry. Antiviral Res 2020;181:104873.

Sheng Cui is the principal investigator of the structural biology group at the Institute of Pathogen Biology, Chinese academy of Medical Sciences. He received his B. S. degree from Wuhan University, and Ph.D. degree from the International School of Advanced Study, Italy. He started independent research in 2009. His research focuses on the molecular and structural basis underlying viral entry and replication using the methods of structural biology, biochemistry, and virology.

Zhongping Tan received his B.S./M.S. degrees from Peking University and Ph.D. degree from Columbia University. After that, he had worked as a postdoctoral research fellow at the Memorial Sloan-Kettering Cancer Center and as an assistant professor at the University of Colorado Boulder. His research focuses on developing a quantitative understanding of the effects of protein glycosylation and providing a guideline for protein glycoengineering.

Wei Hao is an assistant researcher at the Institute of Pathogen Biology, Chinese Academy of Medical Sciences and Peking Union Medical College. He obtained his Ph.D. degree from China Agricultural University, Beijing, under the guidance of Prof. Ji-Jie Chai, in 2014. He mainly works on the structure and function of the key proteins of human pathogens and the molecular mechanisms of interaction between host and microbe.

Bo Ma graduated from Shandong University and received his Ph.D. degree from Peking University. He completed his postdoctoral research at the China Academy of Chinese Medical Sciences. Currently, he is a research assistant professor at the Institute of Materia Medical, Chinese Academy of Medical Science, and Peking Union Medical College. His research focuses on the synthesis of homogeneous glycoproteins and elucidation of the structure-function relationship of protein glycosylation.