

SARS-CoV-2 Spike Protein Interacts with Multiple Innate Immune Receptors

Authors:

Chao Gao¹, Junwei Zeng¹, Nan Jia¹, Kathrin Stavenhagen¹, Yasuyuki Matsumoto¹, Hua Zhang², Jiang Li³, Adam J. Hume^{4,5}, Elke Mühlberger^{4,5}, Irma van Die⁶, Julian Kwan⁷, Kelan Tantisira³, Andrew Emili⁷, and Richard D. Cummings^{1*}

Affiliations:

¹Department of Surgery, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA.

²Laura and Isaac Perlmutter Cancer Center, New York University Langone Medical Center, New York, NY, USA

³Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

⁴Department of Microbiology, Boston University School of Medicine, Boston, MA, USA

⁵National Emerging Infectious Diseases Laboratories, Boston University, Boston, MA, USA

⁶Department of Molecular Cell Biology and Immunology, VU University Medical Center, Amsterdam, The Netherlands

⁷Center for Network Systems Biology, Departments of Biochemistry and Biology, Boston University, Boston, MA, 02118 USA

Email Addresses:

Chao Gao cgao3@bidmc.harvard.edu

Junwei Zeng jzeng1@bidmc.harvard.edu

Nan Jia njia@bidmc.harvard.edu

Kathrin Stavenhagen kstavenh@bidmc.harvard.edu

Yasuyuki Matsumoto ymatsumo@bidmc.harvard.edu

Hua Zhang Hua.Zhang@nyulangone.org

Jiang Li rejia@channing.harvard.edu

Adam J. Hume hume@bu.edu

Elke Mühlberger muehlber@bu.edu

Irma van Die irma.van.die@gmail.com

Julian Kwan jhkwan@bu.edu

Kelan Tantisira kelan.tantisira@channing.harvard.edu

Andrew Emili aemili@bu.edu

* Corresponding author:

Richard D. Cummings, rcummin1@bidmc.harvard.edu

41 **Abstract:**

42 The spike (S) glycoprotein in the envelope of SARS-CoV-2 is densely glycosylated but the
43 functions of its glycosylation are unknown. Here we demonstrate that S is recognized in a
44 glycan-dependent manner by multiple innate immune receptors including the mannose
45 receptor MR/CD206, DC-SIGN/CD209, L-SIGN/CD209L, and MGL/CLEC10A/CD301.
46 Single-cell RNA sequencing analyses indicate that such receptors are highly expressed in
47 innate immune cells in tissues susceptible to SARS-CoV-2 infection. Binding of the above
48 receptors to S is characterized by affinities in the picomolar range and consistent with S
49 glycosylation analysis demonstrating a variety of N- and O-glycans as receptor ligands.
50 These results indicate multiple routes for SARS-CoV-2 to interact with human cells and
51 suggest alternative strategies for therapeutic intervention. (116 words)

52

53 **Introduction:**

54 SARS-CoV-2 is a positive-sense RNA enveloped virus characterized by a surface Spike (S)
55 glycoprotein¹. During host cell invasion, S binds to receptors on cell membranes, such as
56 angiotensin-converting enzyme 2 (ACE2)^{1,2,3}. However, the nature and function of the S
57 glycosylation is not fully understood.

58 Densely glycosylated with multiple Asn-linked (N-glycans) and a few Ser/Thr-linked (O-
59 glycans)^{4,5}, the S protein of SARS-CoV-2 potentially presents ligands for a variety of innate
60 immune receptors, including C-type lectin receptors (CLRs), that are known to bind specific
61 glycans mostly in a Ca²⁺-dependent manner^{6,7}. CLRs such as DC-SIGN/CD209, L-
62 SIGN/CD209L/CLEC4M, mannose receptor/MR/MRC1/CD206, MGL/CLEC10A/CD301,
63 and Dectin-2/CLEC6A, are highly expressed within the human immune system including
64 monocytes, dendritic cells, and macrophages, functioning as the first-line of defense against
65 invading viruses and pathogens^{8,9}. As known pattern-recognition receptors, CLRs, especially
66 DC-SIGN, can direct host immune responses against numerous pathogens in a glycan-
67 specific manner by modulating Toll-like receptor-induced activation¹⁰.

68 Evidence implicates innate immune cells in the pathogenesis of SARS-CoV-2. Over 80% of
69 patients with SARS-CoV-2 infection present with lymphopenia and an increased neutrophil-
70 lymphocyte ratio¹¹. Patients with severe COVID-19 exhibit hyperactive macrophages in the
71 bronchoalveolar lavage fluid (BALF) and oropharyngeal swab¹². Likewise, increased
72 infiltration and activation of macrophages is observed in biopsy or autopsy specimens from
73 COVID-19 patients¹³. Previous studies of the closely-related SARS-CoV demonstrated that
74 primary human monocytes and dendritic cells can be infected^{14,15}, and SARS-CoV S can

75 bind DC-SIGN and L-SIGN^{16, 17, 18, 19}. Eight glycosylation sites of SARS-CoV S were
76 identified to be involved in their interactions^{20, 21}, among which six are conserved in SARS-
77 CoV-2 (**Supplementary Fig. 1**). However, it is not known whether SARS-CoV-2 interacts
78 with a variety of CLRs.

79 Here, we demonstrate that many different CLRs directly bind in a glycan-dependent manner
80 to the S glycoprotein of SARS-CoV-2 with picomolar affinities. Binding of DC-SIGN can
81 trigger the internalization of S in 3T3-DC-SIGN+ cells, which implies the potential
82 involvement in viral entry. Furthermore, we dissect the N- and O-glycan sequences on the
83 recombinant SARS-CoV-2 S, and identify glycan features that are crucial for interactions
84 with these CLRs. The analyses of open accessible single-cell RNA sequencing data confirm
85 that various human tissues and their resident immune cells differentially express CLRs,
86 including MR, MGL and DC-SIGN in bronchoalveolar macrophages in patients with SARS-
87 CoV-2. This is in direct contrast to the absence of ACE2 expression within the same cell
88 types across the tissues. Our study identifies new SARS-CoV-2 binding receptors expressed
89 on innate immune cells, particularly on macrophages and dendritic cells, which could
90 accentuate severe pathological inflammation along with cytokine release syndrome. The
91 results suggest potential additional routes for viral infection and new anti-viral strategies.

92

93 **Results:**

94 **Multiple CLRs bind SARS-CoV-2 S in a glycan-dependent manner**

95 Multiple CLRs including DC-SIGN, L-SIGN, MR (C-type lectin domains 4-7) and MGL
96 exhibited strong binding to the recombinant full-length S produced in human embryonic
97 kidney HEK293 cells (**Fig. 1a-c & e**). HEK293 cells are known to present a spectrum of
98 human glycosylation reflective of the kidney and other epithelial tissues²². DC-SIGN, L-
99 SIGN and MGL also bound to recombinant S1, the subunit involved in ACE2 recognition. By
100 contrast, another CLR, Dectin-2, did not bind S or S1, but bound the positive control, a yeast
101 extract (EBY-100) containing mannan-type ligands (**Fig. 1d**). Binding of DC-SIGN, L-SIGN,
102 MR and MGL was glycan-dependent, as it was sensitive to treatment with sequence-specific
103 glycosidases (**Fig. 1f**). Binding of DC-SIGN and L-SIGN was attenuated by Endoglycosidase
104 H (Endo H, oligomannose and hybrid type N-glycan-targeting) and eliminated by PNGase F
105 (N-glycan-targeting), suggesting that the two CLRs bind to S via both oligomannose and
106 complex N-glycans. MR binding was abolished by both Endo H and PNGase F (**Fig. 1c**),
107 suggesting that oligomannose N-glycans are the ligands. A profound reduction in MGL
108 binding was observed by exposure to PNGase F, indicating that MGL ligands reside

109 primarily on N-glycans. Sialic acid appeared to mask some of the MGL ligands as
110 neuraminidase (Neu) treatment slightly increased the binding (**Fig. 1e**), which was similar to
111 the effect on bovine submaxillary mucin (BSM) with abundant O-glycans comprised of
112 sialylated N-acetylgalactosamine (STn antigen). The effects of glycosidase treatment were
113 confirmed by loss or gain of binding by glycan-binding lectins GNA, ConA and VVA
114 (**Supplementary Fig. 2a, b & d**). Interestingly, neither DC-SIGN nor L-SIGN bound
115 recombinant S2 (**Supplementary Fig. 2e-h**), the subunit mediating cell fusion. In all cases
116 CLR binding to S and S1 was Ca²⁺-dependent as expected (**Supplementary Fig. 2i-l**).
117 Taken together, our results demonstrate that DC-SIGN and L-SIGN bind the recombinant S
118 via high-mannose and complex N-glycans, while MR recognizes S via its high-mannose
119 moieties only. MGL binding can be largely attributed to N-glycans, but O-glycans could also
120 be partly involved in its recognition.

121

122 **CLRs interact with the SARS-CoV-2 S with high affinity**

123 Earlier study showed DC-SIGN and MGL at 1 µg/ml can bind the recombinant receptor
124 binding domain (RBD) of S by ELISA assay²³. Here, using a native-like trimeric SARS-
125 CoV-2 S protein, we sought to measure the binding affinities of DC-SIGN, L-SIGN, MR and
126 MGL, and the canonical SARS-CoV-2 entry receptor ACE2. The results indicate that the
127 CLRs and ACE2 all bind S in a dose-dependent manner (**Fig. 1g-k**). While ACE2 showed the
128 highest affinity with K_d = 4 pM, binding of DC-SIGN, MGL and MR was strong, with K_d =
129 192, 312 and 317 pM, respectively. L-SIGN showed moderate affinity with K_d = 986 pM.

130

131 **SARS-CoV-2 S glycoforms mediate interactions with CLRs**

132 To directly characterize the nature of S glycans that interact with these CLRs, we performed
133 in-depth N-glycan and O-glycopeptide analyses (**Fig. 2**). A similar set of N-glycans,
134 including oligomannose- and complex-type, was detected in the recombinant full-length S,
135 the S1 and S2, but the relative abundance of each individual glycan varies between samples
136 (**Supplementary Fig. 3, Supplementary Table 1**). Oligomannose N-glycans were mainly
137 detected in the full-length S, with the major component being Man₅GlcNAc₂, serving
138 potential ligand for DC-SIGN, L-SIGN, and MR^{24, 25}. Major epitopes on the complex N-
139 glycans were revealed by MALDI-TOF-MS/MS analysis (**Supplementary Fig. 4**). Most
140 complex N-glycans contained core Fuc, and a large proportion were neutral and terminated
141 with either GlcNAc or LacNAc (Gal-GlcNAc) (**Fig. 2**). Notably, our N-glycan microarray

142 analysis revealed that although certain GlcNAc-terminating N-glycans can be bound by DC-
143 SIGN and L-SIGN, the binding was greatly attenuated by the presence of core Fuc (#5 and #7
144 vs #1 and #3, **Supplementary Fig. 5** and **Supplementary Table 2**). Thus, they are unlikely
145 to be the binding partners of DC-SIGN and L-SIGN. Moreover, MS/MS confirmed the
146 presence of Lewis A/X and LacdiNAc epitopes (GalNAc-GlcNAc) (**Supplementary Fig. 4a**
147 **& b**). The former was relatively high in the full-length S and the S1, potentially serving as
148 ligands of DC-SIGN and L-SIGN^{26,27}. The latter was particularly abundant in S1 and S2,
149 potentially supporting the binding of MGL²⁸. The higher molecular-weight region is
150 dominated by multi-antennary N-glycans with various degrees of sialylation
151 (**Supplementary Fig. 3**), which are not known ligands for those CLRs.

152 In an O-glycopeptide-centered analysis of the full-length SARS-CoV-2 S, we identified four
153 new O-glycosylation sites (Tyr28-Arg34, Thr678, Ser 686 and Thr1160) in addition to the
154 previous reported site (Thr323) in the RBD^{4,5} (**Fig. 2**, **Supplementary Fig. 6** and
155 **Supplementary Table 3**). All sites except Thr323 were partially occupied. Major glycan
156 epitopes identified on these sites include non-, mono- and disialylated core 1 (Gal β 1-
157 3GalNAc-R) and core 2 [GlcNAc β 1-6(Gal β 1-3)GalNAc-R]. Importantly, Tn antigen
158 (GalNAc), which is a binding determinant for MGL, was only partially present on Thr323.
159 This is in line with our conclusion that MGL mainly binds N-glycans.

160 Among the four newly discovered O-glycosylation sites, two reside in the furin cleavage site
161 between S1/S2 (Thr678 and Ser686) which are unique to SARS-CoV-2. Although there has
162 not yet been evidence that O-glycosylation plays a role in protease cleavage of S^{1,29}, further
163 investigation is warranted as O-glycosylation does affect protease susceptibility in other
164 systems, as well as antibody recognition^{30,31}.

165 In summary, our glycomics, glycoproteomics analyses, and glycan microarray analysis
166 confirmed that the glycans on SARS-CoV-2 S could serve as ligands for DC-SIGN, L-SIGN,
167 MR and MGL.

168

169 **DC-SIGN and L-SIGN on the cell surface bind SARS-CoV-2 S resulting in its** 170 **internalization**

171 In order to investigate whether S interacts with the CLRs expressed on cell surfaces, we
172 performed flow cytometry using transduced fibroblast-derived 3T3 cells expressing DC-
173 SIGN and L-SIGN, although the latter had only a L-SIGN⁺ subpopulation (**Supplementary**
174 **Fig. 7**). The parental 3T3 cells lacked expression of DC-SIGN or L-SIGN. In accordance

175 with the western blot results, the full-length SARS-CoV-2 S trimer was strongly bound by
176 DC-SIGN⁺ cells (**Fig. 3a & b**) and to a lesser extent by L-SIGN⁺ cells (**Fig. 3c & d**). Ten
177 minutes after incubation, DC-SIGN⁺ cells presented signs of S internalization which was
178 significantly higher than the parental cells after 30 mins (**Fig. 3e**). Thus, our results indicate
179 that SARS-CoV-2 S can be recognized and captured by cells expressing these CLRs, which
180 can lead to internalization of the virus.

181

182 **CLRs bind S produced by SARS-CoV-2-infected Vero E6 cells**

183 To further explore interactions of a native viral S protein with CLRs, we performed western
184 blot using lysates of SARS-CoV-2-infected Vero E6 cells (**Fig. 3f-j**). Twenty-four hours post
185 infection, SARS-CoV-2 S can be robustly detected in the cell lysates by monoclonal antibody
186 1A9 (**Fig. 3f**), which binds the recombinant full-length S but not S1 of SARS-CoV-2
187 (**Supplementary Fig. 7e**). No band was detected in the virus-containing culture supernatant
188 (SN), even at a higher loading amounts (20×).

189 Among the CLRs, DC-SIGN and L-SIGN recognized a band with identical mobility to the
190 full-length S in the lysates of SARS-CoV-2-infected cells (red arrow, **Fig. 3g & h**). This was
191 not present in the mock-infected cells. MR also exhibited positive binding in the SARS-CoV-
192 2-infected cells, and in addition, in the SN (red arrows, **Fig 3i**). By contrast, MGL, although
193 strongly bound to the recombinant S produced in HEK293 cells, did not appear to bind to S in
194 the Vero E6 cell lysates or in the SN (**Fig. 3j**). These results indicate that the glycosylation
195 status of S is influenced by the source of glycosylation, either in infected cells or the
196 recombinant expression system.

197 Taken together, our results demonstrate that SARS-CoV-2-infected Vero E6 cells produce S
198 protein which can be bound by DC-SIGN, L-SIGN and MR, suggesting that the virus could
199 be captured by host cells expressing these CLRs via their unique carbohydrate recognition.

200

201 **CLRs are expressed on innate immune cells in tissues susceptible to SARS-CoV-2** 202 **infection**

203 To characterize the distribution of DC-SIGN, L-SIGN, MR and MGL, we surveyed their
204 expression patterns together with ACE2 in multiple human tissues using public available
205 single-cell RNA sequencing (scRNA-seq) data (**Fig. 4, Supplementary Fig. 8 & 9**).

206 Consistent with previous results³², ACE2 expression was generally low in all analyzed
207 datasets, including human lung and upper airway, thymus, pancreas, spleen, ileum, liver and

208 colon (**Fig. 4a & b** and **Supplementary Fig. 8**). In comparison, expression levels of MR and
209 MGL were higher, particularly in cells of the lung and upper airway (**Fig. 4a**). Their
210 expression was mainly restricted to resident immune cells, such as macrophages and dendritic
211 cells, which was also the major cell types producing DC-SIGN, although to a lower extent. L-
212 SIGN was mainly expressed in endothelial cells in liver and pancreas, in lymphatic tissues in
213 ileum, as well as in immune cells, but at much lower levels (**Fig. 4a & b** and **Supplementary**
214 **Fig. 8**).

215 Emerging evidence showed a strong pathological inflammation in patients with COVID-19,
216 correlated with the presence of monocytes, macrophages and dendritic cells^{33,34}. Therefore,
217 we examined the expression of DC-SIGN, L-SIGN, MR and MGL in available dataset of
218 BALF cells of healthy controls (HC), and patients with moderate (M) and severe (S) COVID-
219 19 (3, 6 and 3 cases, respectively, **Fig. 4c** and **Supplementary Fig. 9**). Consistent with the
220 original report¹², unbiased clustering identified over 30 distinct cell clusters, with the majority
221 of cells being macrophages (**Supplementary Fig. 9a**). Expression of ACE2 was mainly
222 restricted to epithelial cells of severe patients (**Fig. 4c**). In comparison, MR was
223 predominantly expressed in almost all macrophages and dendritic cells from the three donor
224 groups. Of note, the expression levels of DC-SIGN and MGL were increased in severe
225 COVID-19 patients with elevated amount of proinflammatory monocyte-derived
226 macrophages and inflammatory cytokines and chemokines, including IL-6, TNF, CXCL10,
227 CXCL8, IL-1B, CCL2 and CCL3 (**Fig. 4c** and **Supplementary Fig. 9**). The resulting
228 increased systemic cytokine production from activated macrophages may contribute to the
229 pathophysiology of severe COVID-19.

230 Our analyses confirm the expression of CLRs and absence of ACE2 in bronchoalveolar and
231 other innate immune cells, indicating that DC-SIGN, MR and MGL might serve as
232 alternative receptors and entry routes in these cells for SARS-CoV-2.

233

234 **Discussion:**

235 Our results demonstrate that the SARS-CoV-2 S can engage in a glycan-dependent manner
236 with multiple CLRs including DC-SIGN, L-SIGN, MR, and MGL, which are highly
237 expressed in innate immune cells and lymphoid organs, as confirmed by scRNA-seq analysis.
238 The observed interactions with picomolar affinities can initiate receptor-dependent
239 internalization of S as exemplified by DC-SIGN, and potentially provide routes for virus to
240 enter macrophages and dendritic cells. Our scRNA-seq analyses confirmed the co-expression
241 of the CLRs such as DC-SIGN, MR and MGL, and inflammatory cytokines and chemokines

242 in hyperactive macrophages and dendritic cells in patients with COVID-19. This is consistent
243 with the altered cytokine production in SARS-CoV-infected macrophages and dendritic cells,
244 albeit lack productive viral replication^{14, 15}. Our results offer a possible explanation to how
245 SARS-CoV-2 spreads to extrapulmonary tissues within the host^{35, 36, 37}, as the innate immune
246 cells lacking ACE2 expression can still internalize the virus via those CLRs.

247 Glycosylation of SARS-CoV-2 S is obviously complex and dependent on the nature of the
248 protein and the glycosylation machinery of infected host cells. Our results indicate that
249 glycosylation potentially determines the immune receptors with which SARS-CoV-2 interact.
250 This variation may lead to virus clearance or on the contrary, results in spread of the virus to
251 other organs, or even other hosts. Interactions of SARS-CoV-2 with CLRs and subsequent
252 internalization suggests a possibility that resident innate immune cells in the lung may lessen
253 the available titer of free virus by endocytosis, thus slowing the onset of symptoms and
254 contributing to asymptomatic or presymptomatic patients^{38, 39}. Notably, DC-SIGN and L-
255 SIGN can promote “trans” viral transmission, particularly exemplified in HIV-1⁴⁰. Early
256 studies confirmed that DC-SIGN and L-SIGN-expressing cells were able to transfer SARS-
257 CoV to susceptible target cells^{17, 18}. Further research is warranted to explore whether this
258 could occur with SARS-CoV-2. A recent proteomics study discovered strong association
259 between the level of DC-SIGN and variants in the *ABO* locus, the glycosyltransferases
260 required for blood group synthesis, a known genetic risk factor for respiratory failure in
261 COVID-19⁴¹. Given that SARS-CoV-2 RNAemia is detected, particularly in patients with
262 severe disease^{38, 42}, it is tempting to speculate that subsequent to CLR-dependent
263 internalization, SARS-CoV-2 could be conveyed by CLR-expressing innate immune cells,
264 and redistributed to permissive tissues where more damage could occur⁴³. In accord with this
265 hypothesis, a recent study revealed that SARS-CoV-2 nucleocapsid can be detected in
266 spleens and lymph nodes⁴⁴.

267 The results here suggest several diagnostic and therapeutic directions. Inhibition of the CLRs
268 that bind virus or inhibition of the glycan-CLR interactions by small molecules might lessen
269 the distribution of virus in COVID-19 patients and potentially limit immune cell
270 (hyper)activation. Analysis of patients for polymorphisms in the genes encoding CLRs might
271 reveal associations with disease severity, as have been found for L-SIGN in SARS-CoV¹⁹.
272 Finally, defining the differences in glycosylation of S and the variant glycoforms might
273 provide insights into disease severity and spread of the virus in COVID-19 patients.

274

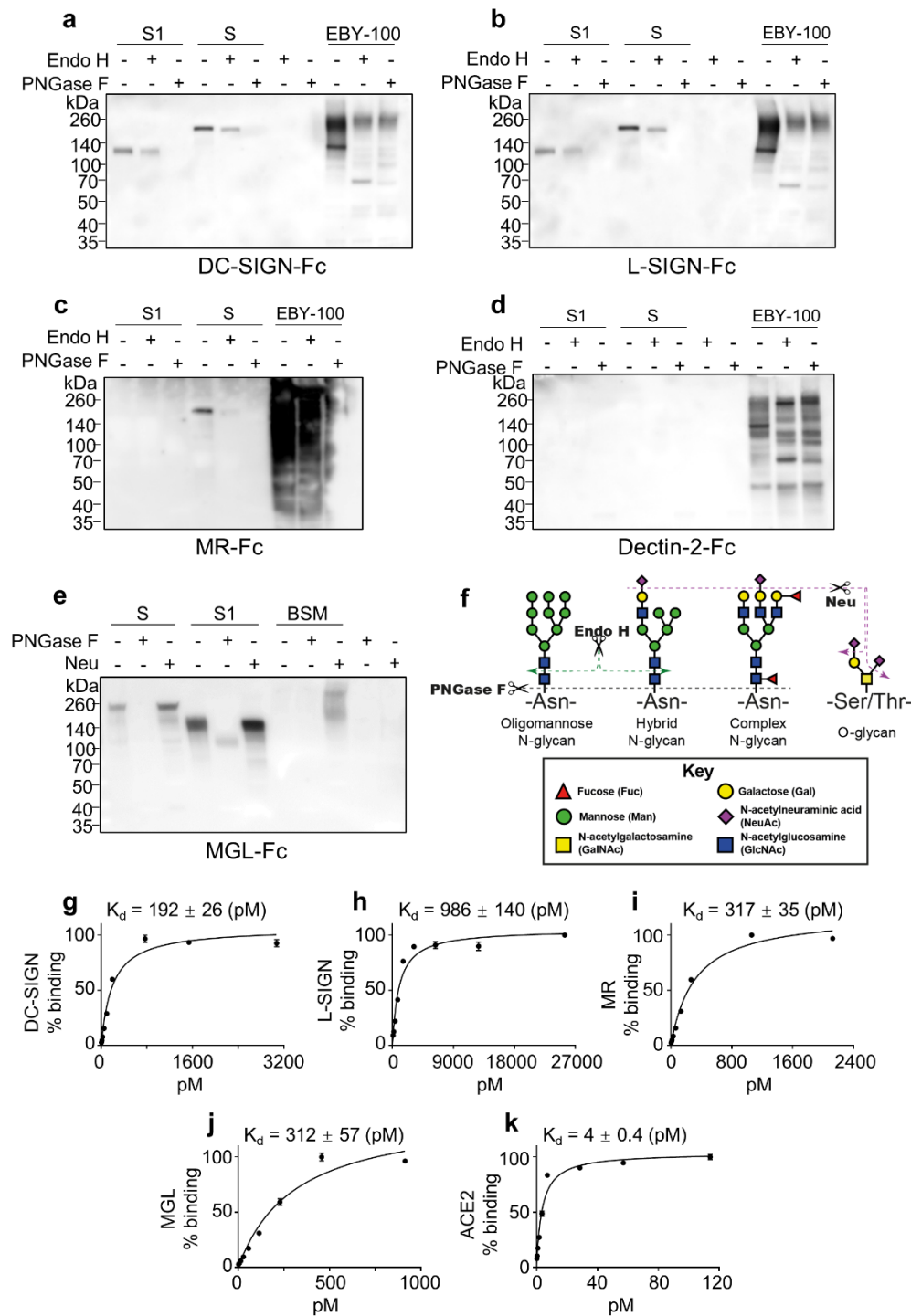
275 **References:**

- 276 1. Hoffmann M, *et al.* SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and
277 Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **181**, 271-280 e278 (2020).
- 278 2. Monteil V, *et al.* Inhibition of SARS-CoV-2 Infections in Engineered Human Tissues
279 Using Clinical-Grade Soluble Human ACE2. *Cell* **181**, 905-913.e907 (2020).
- 280 3. Letko M, Marzi A, Munster V. Functional assessment of cell entry and receptor usage
281 for SARS-CoV-2 and other lineage B betacoronaviruses. *Nat Microbiol* **5**, 562-569
282 (2020).
- 283 4. Shajahan A, Supekar NT, Gleinich AS, Azadi P. Deducing the N- and O-
284 glycosylation profile of the spike protein of novel coronavirus SARS-CoV-2.
285 *Glycobiology*, 2020.2004.2001.020966 (2020).
- 286 5. Watanabe Y, Allen JD, Wrapp D, McLellan JS, Crispin M. Site-specific glycan
287 analysis of the SARS-CoV-2 spike. *Science*, 2020.2003.2026.010322 (2020).
- 288 6. Drouin M, Saenz J, Chiffolleau E. C-Type Lectin-Like Receptors: Head or Tail in Cell
289 Death Immunity. *Front Immunol* **11**, 251 (2020).
- 290 7. van Kooyk Y. C-type lectins on dendritic cells: key modulators for the induction of
291 immune responses. *Biochem Soc Trans* **36**, 1478-1481 (2008).
- 292 8. van Kooyk Y, Rabinovich GA. Protein-glycan interactions in the control of innate and
293 adaptive immune responses. *Nat Immunol* **9**, 593-601 (2008).
- 294 9. Routhu NK, *et al.* Glycosylation of Zika Virus is Important in Host-Virus Interaction
295 and Pathogenic Potential. *International journal of molecular sciences* **20**, 5206
296 (2019).
- 297 10. Gringhuis SI, den Dunnen J, Litjens M, van der Vlist M, Geijtenbeek TBH.
298 Carbohydrate-specific signaling through the DC-SIGN signalosome tailors immunity
299 to Mycobacterium tuberculosis, HIV-1 and Helicobacter pylori. *Nature Immunology*
300 **10**, 1081-1088 (2009).
- 301 11. Guan W-j, *et al.* Clinical Characteristics of Coronavirus Disease 2019 in China. *New*
302 *England Journal of Medicine* **382**, 1708-1720 (2020).
- 303 12. Liao M, *et al.* Single-cell landscape of bronchoalveolar immune cells in patients with
304 COVID-19. *Nature Medicine*, (2020).
- 305 13. Tian S, Hu W, Niu L, Liu H, Xu H, Xiao S-Y. Pulmonary Pathology of Early-Phase
306 2019 Novel Coronavirus (COVID-19) Pneumonia in Two Patients With Lung Cancer.
307 *Journal of Thoracic Oncology* **15**, 700-704 (2020).
- 308 14. Law HKW, *et al.* Chemokine up-regulation in SARS-coronavirus-infected,
309 monocyte-derived human dendritic cells. *Blood* **106**, 2366-2374 (2005).
- 310 15. Tseng C-TK, Perrone LA, Zhu H, Makino S, Peters CJ. Severe Acute Respiratory
311 Syndrome and the Innate Immune Responses: Modulation of Effector Cell Function
312 without Productive Infection. *The Journal of Immunology* **174**, 7977-7985 (2005).
- 313 16. Jeffers SA, *et al.* CD209L (L-SIGN) is a receptor for severe acute respiratory
314 syndrome coronavirus. *Proceedings of the National Academy of Sciences of the*
315 *United States of America* **101**, 15748-15753 (2004).

- 316 17. Marzi A, *et al.* DC-SIGN and DC-SIGNR Interact with the Glycoprotein of Marburg
317 Virus and the S Protein of Severe Acute Respiratory Syndrome Coronavirus. *Journal*
318 *of Virology* **78**, 12090-12095 (2004).
- 319 18. Yang Z-Y, *et al.* pH-Dependent Entry of Severe Acute Respiratory Syndrome
320 Coronavirus Is Mediated by the Spike Glycoprotein and Enhanced by Dendritic Cell
321 Transfer through DC-SIGN. *Journal of Virology* **78**, 5642-5650 (2004).
- 322 19. Chan VSF, *et al.* Homozygous L-SIGN (CLEC4M) plays a protective role in SARS
323 coronavirus infection. *Nature Genetics* **38**, 38-46 (2006).
- 324 20. Shih Y-P, *et al.* Identifying Epitopes Responsible for Neutralizing Antibody and DC-
325 SIGN Binding on the Spike Glycoprotein of the Severe Acute Respiratory Syndrome
326 Coronavirus. *Journal of Virology* **80**, 10315-10324 (2006).
- 327 21. Han DP, Lohani M, Cho MW. Specific Asparagine-Linked Glycosylation Sites Are
328 Critical for DC-SIGN- and L-SIGN-Mediated Severe Acute Respiratory Syndrome
329 Coronavirus Entry. *Journal of Virology* **81**, 12029-12039 (2007).
- 330 22. Goh JB, Ng SK. Impact of host cell line choice on glycan profile. *Crit Rev Biotechnol*
331 **38**, 851-867 (2018).
- 332 23. Chiodo F, *et al.* Novel ACE2-Independent Carbohydrate-Binding of SARS-CoV-2
333 Spike Protein to Host Lectins and Lung Microbiota. *bioRxiv*, 2020.2005.2013.092478
334 (2020).
- 335 24. Guo Y, *et al.* Structural basis for distinct ligand-binding and targeting properties of
336 the receptors DC-SIGN and DC-SIGNR. *Nat Struct Mol Biol* **11**, 591-598 (2004).
- 337 25. Martinez-Pomares L. The mannose receptor. *Journal of Leukocyte Biology* **92**, 1177-
338 1186 (2012).
- 339 26. Feinberg H, Mitchell DA, Drickamer K, Weis WI. Structural Basis for Selective
340 Recognition of Oligosaccharides by DC-SIGN and DC-SIGNR. *Science* **294**, 2163-
341 2166 (2001).
- 342 27. Noll AJ, *et al.* Human DC-SIGN Binds Specific Human Milk Glycans. *Biochemical*
343 *Journal*, (2016).
- 344 28. Mortezaei N, *et al.* Tumor-associated Neu5Ac-Tn and Neu5Gc-Tn antigens bind to C-
345 type lectin CLEC10A (CD301, MGL). *Glycobiology* **23**, 844-852 (2013).
- 346 29. Walls AC, Park Y-J, Tortorici MA, Wall A, McGuire AT, Velesler D. Structure,
347 Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **181**, 281-
348 292.e286 (2020).
- 349 30. de las Rivas M, *et al.* Molecular basis for fibroblast growth factor 23 O-glycosylation
350 by GalNAc-T3. *Nature Chemical Biology* **16**, 351-360 (2020).
- 351 31. Silver ZA, *et al.* Discovery of O-Linked Carbohydrate on HIV-1 Envelope and Its
352 Role in Shielding against One Category of Broadly Neutralizing Antibodies. *Cell*
353 *Reports* **30**, 1862-1869.e1864 (2020).
- 354 32. Sungnak W, *et al.* SARS-CoV-2 entry factors are highly expressed in nasal epithelial
355 cells together with innate immune genes. *Nature Medicine* **26**, 681-687 (2020).

- 356 33. Merad M, Martin JC. Pathological inflammation in patients with COVID-19: a key
357 role for monocytes and macrophages. *Nature Reviews Immunology* **20**, 355-362
358 (2020).
- 359 34. Vabret N, *et al.* Immunology of COVID-19: current state of the science. *Immunity*,
360 (2020).
- 361 35. Wang W, *et al.* Detection of SARS-CoV-2 in Different Types of Clinical Specimens.
362 *JAMA* **323**, 1843-1844 (2020).
- 363 36. Wu Y, *et al.* Prolonged presence of SARS-CoV-2 viral RNA in faecal samples. *The*
364 *Lancet Gastroenterology & Hepatology* **5**, 434-435 (2020).
- 365 37. Lamers MM, *et al.* SARS-CoV-2 productively infects human gut enterocytes.
366 *Science*, eabc1669 (2020).
- 367 38. Huang C, *et al.* Clinical features of patients infected with 2019 novel coronavirus in
368 Wuhan, China. *The Lancet* **395**, 497-506 (2020).
- 369 39. Arons MM, *et al.* Presymptomatic SARS-CoV-2 Infections and Transmission in a
370 Skilled Nursing Facility. *New England Journal of Medicine* **382**, 2081-2090 (2020).
- 371 40. Geijtenbeek TBH, *et al.* DC-SIGN, a Dendritic Cell-Specific HIV-1-Binding Protein
372 that Enhances trans-Infection of T Cells. *Cell* **100**, 587-597 (2000).
- 373 41. Brufsky A, Lotze MT. DC/L-SIGNs of Hope in the COVID-19 Pandemic. *Journal of*
374 *Medical Virology* **n/a**.
- 375 42. Hogan CA, *et al.* High frequency of SARS-CoV-2 RNAemia and association with
376 severe disease. *medRxiv*, 2020.2004.2026.20080101 (2020).
- 377 43. Li H, *et al.* SARS-CoV-2 and viral sepsis: observations and hypotheses. *The Lancet*
378 **395**, 1517-1520 (2020).
- 379 44. chen y, *et al.* The Novel Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-
380 CoV-2) Directly Decimates Human Spleens and Lymph Nodes. *medRxiv*,
381 2020.2003.2027.20045427 (2020).
- 382 45. Wu L, Martin TD, Vazeux R, Unutmaz D, KewalRamani VN. Functional Evaluation
383 of DC-SIGN Monoclonal Antibodies Reveals DC-SIGN Interactions with ICAM-3
384 Do Not Promote Human Immunodeficiency Virus Type 1 Transmission. *Journal of*
385 *virology* **76**, 5905-5914 (2002).
- 386 46. Matsumoto Y, *et al.* Identification of Tn antigen O-GalNAc-expressing glycoproteins
387 in human carcinomas using novel anti-Tn recombinant antibodies. *Glycobiology* **30**,
388 282-300 (2019).
- 389 47. Gao C, *et al.* Unique Binding Specificities of Proteins toward Isomeric Asparagine-
390 Linked Glycans. *Cell chemical biology* **26**, 535-547 e534 (2019).
- 391 48. Jia N, *et al.* The Human Lung Glycome Reveals Novel Glycan Ligands for Influenza
392 A Virus. *Scientific Reports* **10**, 5320 (2020).
- 393 49. Plomp R, *et al.* Site-Specific N-Glycosylation Analysis of Human Immunoglobulin E.
394 *Journal of proteome research* **13**, 536-546 (2014).
- 395 50. Deprez M, *et al.* A single-cell atlas of the human healthy airways. *bioRxiv*,
396 2019.2012.2021.884759 (2019).

- 397 51. Vieira Braga FA, *et al.* A cellular census of human lungs identifies novel cell states in
398 health and in asthma. *Nature Medicine* **25**, 1153-1163 (2019).
- 399 52. Martin JC, *et al.* Single-Cell Analysis of Crohn's Disease Lesions Identifies
400 a Pathogenic Cellular Module Associated with Resistance to Anti-TNF Therapy. *Cell*
401 **178**, 1493-1508.e1420 (2019).
- 402 53. Smillie CS, *et al.* Intra- and Inter-cellular Rewiring of the Human Colon during
403 Ulcerative Colitis. *Cell* **178**, 714-730.e722 (2019).
- 404 54. Baron M, *et al.* A Single-Cell Transcriptomic Map of the Human and Mouse Pancreas
405 Reveals Inter- and Intra-cell Population Structure. *Cell Systems* **3**, 346-360.e344
406 (2016).
- 407 55. Madisson E, *et al.* scRNA-seq assessment of the human lung, spleen, and esophagus
408 tissue stability after cold preservation. *Genome Biology* **21**, 1 (2019).
- 409 56. Park J-E, *et al.* A cell atlas of human thymic development defines T cell repertoire
410 formation. *Science* **367**, eaay3224 (2020).
- 411 57. Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data
412 analysis. *Genome Biol* **19**, 15 (2018).
- 413 58. Stuart T, *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902
414 e1821 (2019).
- 415 59. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell
416 transcriptomic data across different conditions, technologies, and species. *Nat*
417 *Biotechnol* **36**, 411-420 (2018).
- 418
419
420



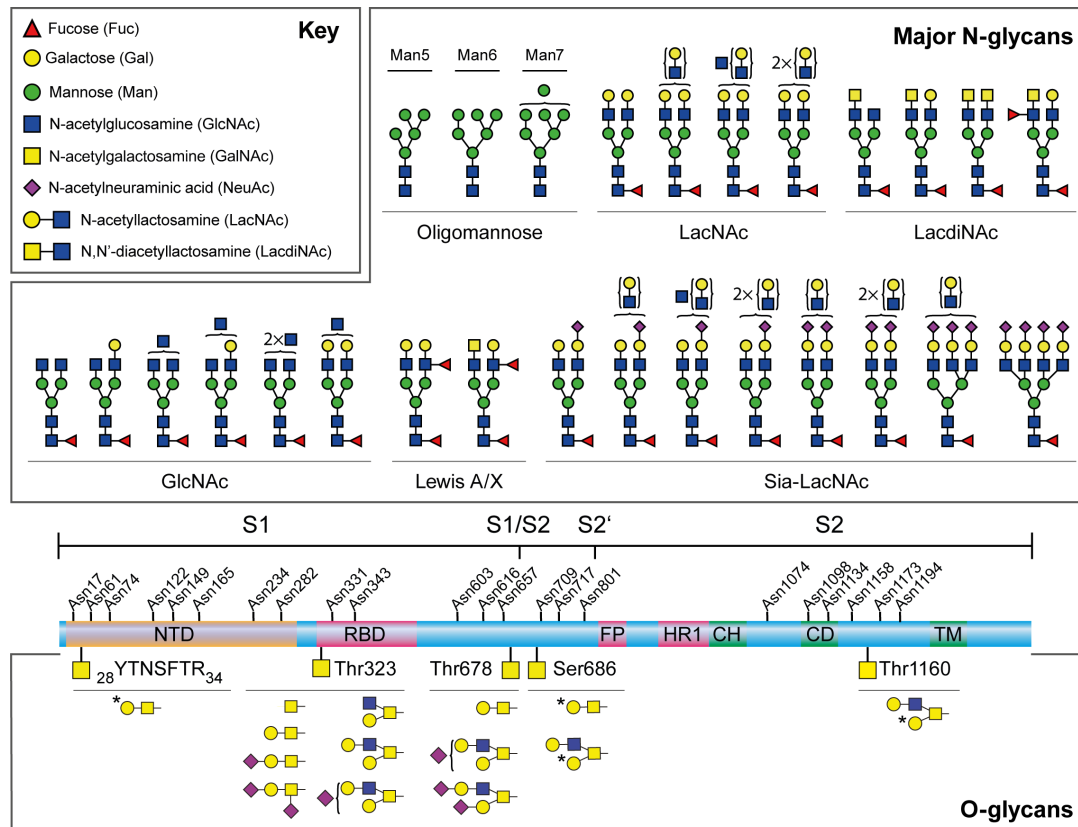
421

422 **Fig. 1. Binding of CLR receptors to the recombinant SARS-CoV-2 S.** (a-e) Immunoblots with
 423 human Fc-fused CLR receptors DC-SIGN (a), L-SIGN (b), MR (c), Dectin-2 (d) and MGL (e)
 424 to detect recombinant S1 and S after mock enzymatic digestion or with Endo H, PNGase F or
 425 Neu digestion. As negative controls, these glycosidases were also included in some assays.
 426 EBY-100 represents the lysates of yeast strain EBY-100. BSM is the recombinant bovine
 427 submaxillary mucin. In all assays 5 mM Ca^{2+} was included in solutions of CLR receptors.
 428 (f) Schematic presentation of the cleavage sites of Endo H, PNGase F and Neu on N- and O-

429 glycans. Endo H cleaves the oligomannose and hybrid N-glycans, while PNGase F removes
430 all N-glycans including the complex type. Neu removes all sialic acids on N- or O-glycans.
431 **(g-k)** Affinity constant measurement for DC-SIGN **(g)**, L-SIGN **(h)**, MR **(i)**, MGL **(j)** and
432 ACE2 **(k)** by ELISA assay. The plates were coated by recombinant SARS-CoV-2 S trimer.
433 Error bars represent SD of two replicates. The data were plotted as % binding relative to the
434 saturated binding as 100%. In all assays 5 mM Ca²⁺ and Tween-20 were included in solutions
435 of CLRs.

436

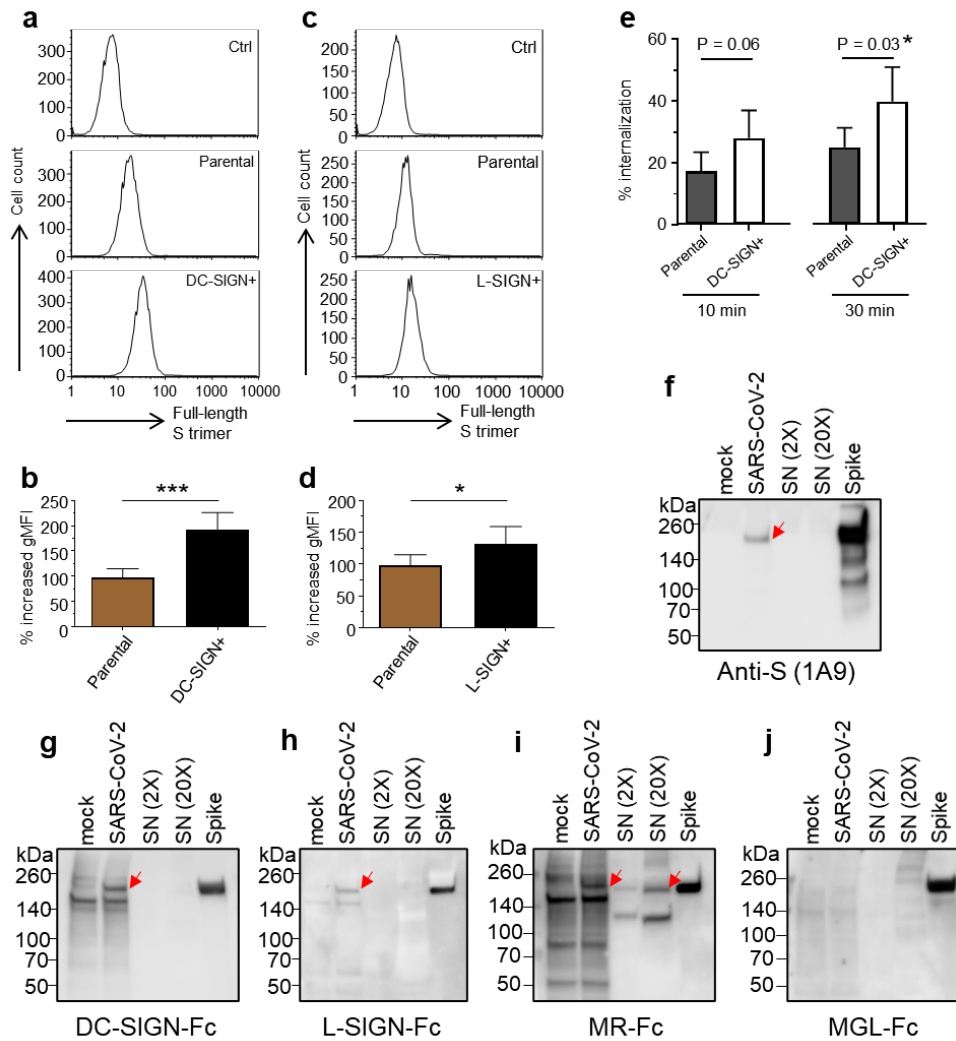
437



438

439

440 **Fig. 2. Major N-glycans and O-glycopeptides identified in the recombinant full-length**
 441 **SARS-CoV-2 S.** Schematic representation of SARS-CoV-2 S shown in the middle. The
 442 positions of N-glycosylation sites are shown on top. Protein domains in the illustration are:
 443 N-terminal domain (NTD), receptor-binding domain (RBD), fusion peptide (FP), heptad
 444 repeat 1 (HR1), central helix (CH), connector domain (CD), and transmembrane
 445 domain (TM). The cleavage sites of S1/S2 and S2' are labelled. Major N-glycan structures detected
 446 by mass spectrometry were categorized by their epitopes on the non-reducing terminal and
 447 shown on top. Cartoon symbols above a curly parenthesis indicates sequences corresponding
 448 to these compositions cannot be unequivocally defined. The structures presented are only the
 449 major glycans on the recombinant full-length S, S1, and S2. A full list of glycans can be
 450 found in **Supplementary Table 1**. The O-glycopeptides detected in the full-length S protein
 451 are presented at the bottom. The identified O-glycosylation sites are marked on the protein
 452 and the O-glycans on each specific site are listed below each site. The LC-MS/MS spectrum
 453 of each O-glycopeptide can be found in **Supplementary Fig. 6**. Labelled with asterisk were
 454 those found only after neuraminidase treatment.
 455

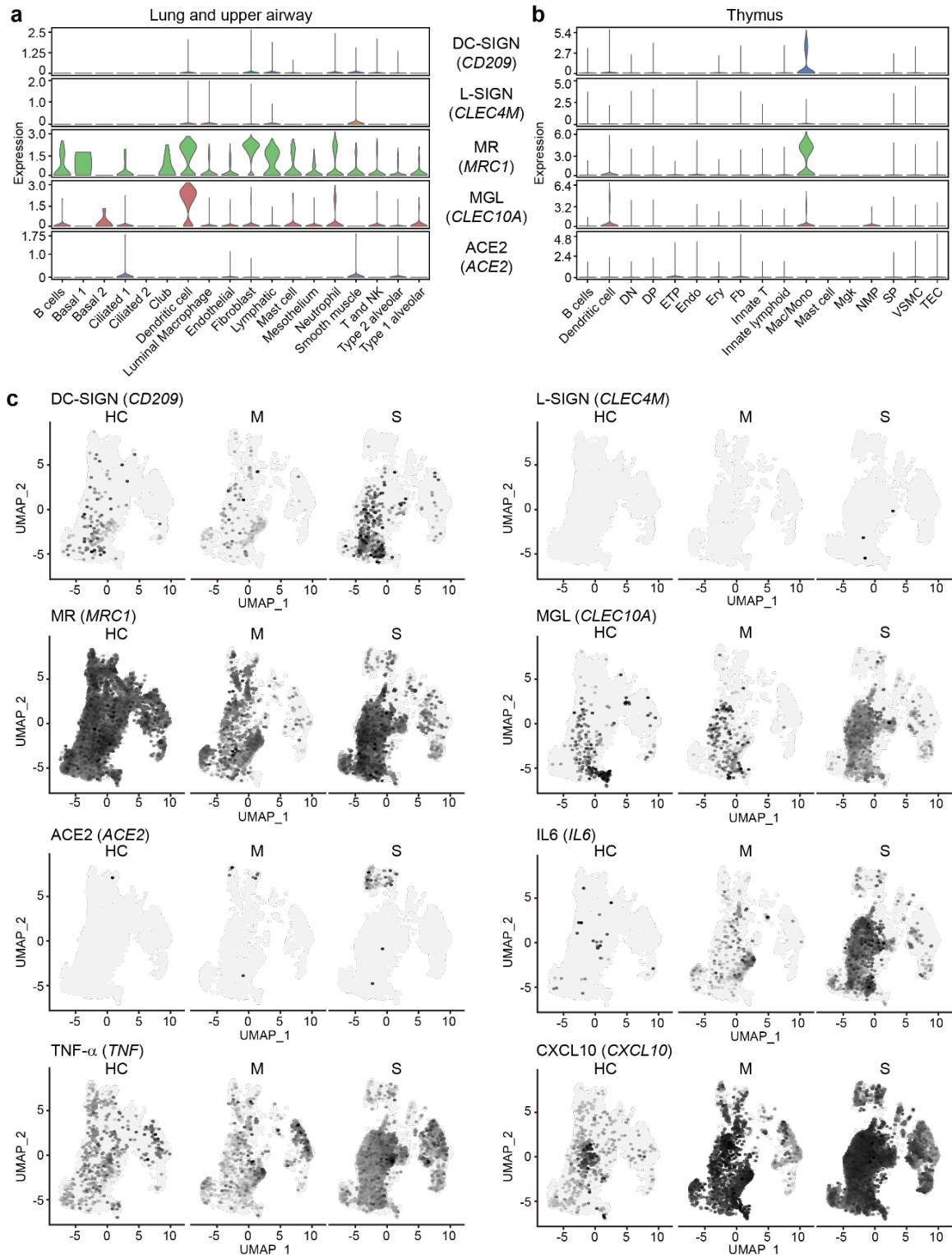


456

457

458 **Fig. 3. Binding analysis of cells expressing DC-SIGN and L-SIGN and Vero E6 cells**
 459 **infected by SARS-CoV-2. (a and c)** Flow cytometry profiles showing the binding of the full-
 460 full-length SARS-CoV-2 S trimer to parental 3T3 cells (middle), and 3T3-DC-SIGN+ and 3T3-L-
 461 SIGN+ cells (bottom). **(b and d)** Increased binding of SARS-CoV-2 S trimer to the parental
 462 3T3 cells, the 3T3-DC-SIGN+ and 3T3-L-SIGN+ cells relative to the secondary antibody
 463 control. Data presented here is the percentage of increase in geometric mean fluorescence
 464 intensity (gMFI). **(e)** Internalization of SARS-CoV-2 S trimer in 3T3-DC-SIGN+ cells
 465 compared to the parental 3T3 cells at 10 and 30 min. **(f-j)** Immunoblots with anti-S mAb 1A9
 466 **(f)**, human Fc-fused CLR proteins DC-SIGN **(g)**, L-SIGN **(h)**, MR **(i)**, and MGL **(j)** using lysates of
 467 Vero E6 cells mock infected (lane 1) or SARS-CoV-2 infected (lane 2), and the SARS-CoV-
 468 2-containing culture supernatant (2X and 20X in lane 3 and 4, respectively). Red arrow
 469 indicates the bands corresponding to the SARS-CoV-2 S. In all assays 5 mM Ca²⁺ was
 470 included in solutions of CLR proteins. The recombinant SARS-CoV-2 S was used as a positive
 471 control.

472



473

474

475 **Fig. 4. Expression of CLR genes in human tissues and CLR/chemokines/chemokines in BALF**

476 **from COVID-19 patients. (a and b) Single-cell transcriptomic analysis of CLR genes**

477 **expression (DC-SIGN, L-SIGN, MR, MGL) and ACE2 in lung and upper airway and thymus**

478 **as indicated. NK, natural killer cells; DN, double-negative T cells; DP, double-positive T**

479 **cells; ETP, early thymic progenitor; Endo, endothelial cells; Ery, erythrocytes; Fb,**

480 fibroblasts; Mono, monocyte; Mac, macrophage; Mgc, megakaryocyte; NMP, neutrophil-
481 myeloid progenitor; SP, single-positive T cells; VSMC, vascular smooth muscle cells; TEC,
482 thymic epithelial cells). (c) UMAP showing the gene expression levels of CLRs gene
483 expression (DC-SIGN, L-SIGN, MR, MGL) and ACE2, and selected cytokines and
484 chemokines in BALF immune cells from health controls (HC, n = 3), moderate cases (M, n =
485 3) and severe/critical cases (S, n = 6).

486

487

488 **Methods:**

489 **Recombinant proteins**

490 Recombinant S1 (His-tag), S2 (His-tag), full-length S protein (with mutations R683A and
491 R685A, His-tag), S protein trimer (with mutations R683A and R685A, His-tag), human
492 ACE2 (Fc-tag) were all purchased from Acrobiosystems (S1N-C52H3, S2N-C52H5, SPN-
493 C52H4, SPN-C52H8, AC2-H5257, respectively). The trimeric conformation of the full-
494 length S protein was validated by the vender using size exclusion chromatography under
495 detection with multi angle light scattering. Human DC-SIGN (Fc-tag), DC-SIGNR (Fc-tag),
496 MGL (Fc-tag) were purchased from Sino Biological (10200-H01H, 10559-H01H, 10821-
497 H01H, respectively). Anti-Spike protein antibody (1A9), which was a mouse monoclonal
498 antibody (IgG1) detecting the spike proteins of both SARS-CoV and SARS-CoV-2 through
499 S2 subunit, was purchased from GeneTex (GTX632604).

500 Human IgG Fc-tagged MR (MR-Fc), containing the murine C-type lectin domains 4-7 fused
501 to a human IgG Fc-portion, was produced in house. HEK293T cells were transfected with
502 MR-Fc DNA (kind gift from L. Martinez Pomares). Transfection was performed with
503 Lipofectamine LTX with Plus Reagent (Thermo Fisher Scientific) conform with the
504 manufacturer's guidelines. The cells were incubated with Lipofectamine complex at 37°C in
505 a CO₂ incubator for 24 h, and the medium refreshed after 24 h. After 9 days the supernatant
506 containing the produced MR-Fc was collected and purified. MR-Fc was purified from cell
507 culture supernatant using Protein A-agarose beads (Roche). For ELISA experiments, the MR-
508 Fc was purified over mannan-agarose beads (Sigma) and protein A-agarose beads. MR-Fc
509 was quantified by Nanodrop (Thermo Fisher Scientific) and stored at -20°C until further use.

510

511 **Cell culture**

512 Human fibroblast cell line 3T3, and the DC-SIGN-, and L-SIGN-transduced 3T3 cells (3T3-
513 DC-SIGN+ and 3T3-DC-SIGNR+, respectively) were obtained through the AIDS Reagent
514 Program, Division of AIDS, NIAID, NIH from Drs. Thomas D. Martin and Vineet N.
515 KewalRamani, HIV Drug Resistance Program, NCI⁴⁵. They were cultured in Dulbecco's
516 Modified Eagle's Medium (DMEM) supplemented with 10% (vol/vol) fetal bovine serum
517 (FBS), 1% penicillin-streptomycin and 2 mM glutamine at 37°C in 5% CO₂. Vero E6 cells
518 (ATCC CRL-1586) were cultured in DMEM supplemented with 2 mM L-glutamine, 50
519 units/ml penicillin, 50 mg/ml streptomycin, and 10% FBS. Infection studies were performed
520 in cell culture medium supplemented with 2% FBS.

521 The *Saccharomyces* strain EBY-100 was purchased from ATCC and cultured under the
522 recommended condition by the vendor. The EBY-100 cell extracts were prepared with the
523 lysis buffer (20 mM Tris, 100 mM NaCl, 1 mM EDTA, 2% Triton X-100, 1% SDS, pH 7.6).
524 The protein concentration was quantified by Pierce™ BCA Protein Assay Kit (Thermo Fisher
525 Scientific).

526

527 **Preparation of SARS-CoV-2**

528 All work with infectious SARS-CoV-2, isolate USA_WA1/2020 was performed at the
529 National Emerging Infectious Diseases Laboratories, Boston University under Biosafety
530 Level 4. Vero E6 cells seeded in 6-well plates were infected with SARS-CoV-2 at a
531 multiplicity of infection (MOI) of 1 or mock infected. After 1 h of virus adsorption at 37°C in
532 5% CO₂, the inocula were removed and replaced with DMEM supplemented with 2% FBS.
533 Twenty-four hours post-infection, cell supernatants were removed, the cells were scraped and
534 pelleted by low-speed centrifugation. Cell pellets were washed once with PBS and
535 resuspended in 270 µl Cell Extraction Buffer (Thermo Fisher Scientific). A 900 µl aliquot of
536 a SARS-CoV-2 stock (Titer: 1.58 x10⁷ TCID₅₀ units/ml) was also used for analysis. Both the
537 cell lysates and the viral stock were inactivated by addition of SDS (1 % final concentration)
538 and boiling for 10 minutes prior to removal from the Biosafety Level 4 facility. The
539 inactivated cell lysates and virus stock were aliquoted and stored at -80°C until further use.
540 The protein concentration was quantified by Pierce™ BCA Protein Assay Kit (Thermo Fisher
541 Scientific).

542

543 **Western Blot and Lectin Blot**

544 The recombinant S, S1 and S2 proteins, together with controls were subject to enzymatic
545 digestion by PNGase F (P0708L, New England Biolabs), Endo H (P0702L, New England
546 Biolabs), neuraminidase (10269611001, Roche) or mock treatment before loading onto the
547 SDS-PAGE gel. The enzymatic digestion was following the protocols recommended by the
548 vendors. Recombinant S, S1 and S2 protein and BSM were used at 1µg/lane, while the EBY-
549 100 extract was used at 15 µg/lane. The reaction mixture was directly loaded on the 4-20%
550 SDS-PAGE gel (GenScript) following addition of 4× laemmli loading buffer. The cell
551 extracts either from SARS-CoV-2 infected or mock infected Vero E6 cells (both at 50 µg
552 protein/lane), together with the culture supernatant (at 12 µg (2×) and 120 µg (20×)
553 protein/lane), were loaded on the SDS-PAGE gel following addition of 4× laemmli loading

554 buffer and 10 min heating. The gels were either stained with Coomassie Brilliant Blue or
555 Colloidal Blue (both from Thermo Fisher Scientific) to visualize the proteins, or transferred
556 to a PVDF membrane (Thermo Fisher Scientific) for Western blots.

557

558 The Western blot and lectin blot analysis were performed following the protocols published
559 previously⁴⁶. The proteins were under the following concentrations: DC-SIGN, L-SIGN,
560 Dectin-2, MGL all at 1 µg/ml, MR at 2.5 µg/ml, biotinylated plant lectins GNA and VVA
561 both at 1 µg/ml, Con A at 0.1 µg/ml, or antibodies mAb100 at 1 µg/ml and anti-spike
562 antibody 1A9 at 0.5 µg/ml. The CLRs were prepared either in BSA-TTBS buffer [1% BSA in
563 20 mM Tris, 300 mM NaCl, 5 mM CaCl₂, 2 mM MgCl₂, pH 7.4, with 0.05% Tween-20] or
564 BSA-TTBS-EDTA [1% BSA in 20 mM Tris, 300 mM NaCl, 10 mM EDTA, pH 7.4] with
565 0.05% Tween-20 (TTBS)] to check the calcium dependence. HRP-labelled secondary
566 reagents including goat anti-human IgG-HRP, goat anti-mouse IgG-HRP (Jackson
567 ImmunoResearch Laboratories) or streptavidin-HRP (Vector Laboratories) are all used in
568 1:10,000 dilution.

569

570 **ELISA**

571 The ELISA assay was performed following the protocols published previously⁴⁶. The S-
572 protein trimer was immobilized in a 96-well plate (0.2 µg/well) overnight at 4°C. The
573 proteins were serial diluted (1:2) and applied to each well. HRP-labeled goat anti-human IgG
574 was used at 1:1,000 dilution and the color development was with TMB ELISA Substrate
575 (Abcam, ab171523) The experiment was performed in duplicate and the background
576 subtraction was done by subtracting the corresponding binding signals obtained with BSA-
577 immobilized wells. Affinity constant was calculated with GraphPad Prism 6.0 (GraphPad
578 Software, Inc.).

579

580 **Flow cytometry**

581 For cell surface binding assay, the cultured 3T3, 3T3-DC-SIGN and 3T3-L-SIGN cells were
582 collected and washed with cold PBS once. The cells were incubated with full-length S protein
583 trimer (20 µg/ml) or the monoclonal antibody against DC-SIGN, #120507 (2 µg/ml) and L-
584 SIGN, #120604 (2 µg/ml), or buffer only (negative control) on ice for 30 minutes. Binding of
585 the full-length S protein was detected by mouse anti-His IgG (Invitrogen MA121315, 2
586 µg/ml) and Alexa Fluor-488-labelled goat anti-mouse IgG (Invitrogen A11001, 1:500).

587 Binding of anti-DC-SIGN and anti-L-SIGN were detected directly with Alexa Fluor 488-
588 labelled goat anti-mouse IgG. The results shown were from three independent experiments
589 analyzed by FlowJo software.

590

591 **Internalization assay**

592 To measure internalization, the cells were suspended and incubated with S protein trimer on
593 ice as stated above. The cells were then centrifuged and fixed in 0.5% paraformaldehyde (T0)
594 in binding buffer or allowed for further incubation at 37 °C for 10 or 30 min (Tn) and ended
595 with 0.5% paraformaldehyde. Fixed cells were stained with mouse anti-His IgG and Alexa
596 Fluor-488-labelled goat anti-mouse IgG as indicated above for residual S trimer detection.
597 Internalization rate (%) was calculated by the formula: [gMFI of S trimer(T0)-gMFI of S
598 trimer (Tn)]/gMFI of S trimer (T0) × 100. The results shown were from five independent
599 experiments with comparable results.

600

601 **Microarray analysis**

602 Microarray analyses on the complex N-glycan array were conducted as reported previously⁴⁷.
603 The microarray slides were probed with Fc-tagged DC-SIGN and L-SIGN at 10 µg/ml
604 diluted in 1% BSA (TSM with 0.05% Tween) and binding was detected with Alexa Fluor 488
605 labelled goat anti-human IgG (H+L) (Invitrogen) at 5 µg/ml.

606

607 **Glycomics analysis**

608 The full length S, S1 and S2 proteins were subjected to N-glycomics analysis following the
609 protocols published previously⁴⁸ with some modifications. Briefly, 20 µg each of the
610 recombinant proteins were lyophilized and digested with PNGase F (P0701S, New England
611 Biolabs Inc.) according to the manufacture's instruction. Post digestion, the glycans were
612 purified by Sep-Pak C18 cartridges (WAT054955, Waters Corp.) and subjected to
613 permethylation. The permethylated glycans were purified by chloroform extraction and Sep-
614 Pak C18 cartridges prior to mass spectrometric analysis.

615 MALDI-TOF MS and MALDI-TOF/TOF MS/MS analysis of the permethylated glycans
616 were performed with the UltrafleXreme mass spectrometer (Bruker Corp.) equipped with a
617 Smartbeam II laser. Spectrum between mass-to-charge (m/z) 1000 and 5500 was acquired

618 under reflectron positive mode. Selective peaks were subject to MS/MS analysis. Each MS
619 spectrum presented an accumulation of 20,000 laser shots.

620

621 **O-glycan LC-MS analysis**

622 Three microgram of full-length S protein were treated with PNGase F (NEB) for N-glycan
623 release according to the manufacture's instruction, followed by optional sialic acid removal
624 using 0.02 U sialidase (Roche), prior to in-gel trypsin treatment⁴⁹. After overnight trypsin
625 treatment an optional AspN (Roche) digestion was performed in 1:25 ratio (enzyme:protein).
626 The samples were dried down in a speed vac concentrator, reconstituted in Cal and further
627 diluted in 0.1 % FA for subsequent LC-MS analysis. All samples were prepared in triplicates.

628 LC-MS was performed using an Ultimate 3000 nano LC coupled to an Orbitrap Fusion
629 Lumos mass spectrometer (both from Thermo Fisher). Three microliters of each sample were
630 loaded onto a C18 precolumn (C18 PepMap 100, 300 μm x 5 mm, 5 μm , 100 \AA , Thermo
631 Fisher Scientific) with 15 $\mu\text{L}/\text{min}$ solvent A (0.1% FA in H_2O) for 3 min and separated on a
632 C18 analytical column (picofrit 75 μm ID x 150 mm, 3 μm , New Objective) using a linear
633 gradient of 2 % to 45 % solvent B (80% acetonitrile, 0.1% FA) over 106 min at 400 nL/min.
634 The ion source was set at 2100 V spray voltage and 200 $^\circ\text{C}$ ion transfer tube temperature. MS
635 scans were performed in the orbitrap at a resolution of 60000 within a scan range of m/z 600
636 – m/z 1600, a RF lens of 30%, AGC target of $1e5$ for a maximum injection time of 50 ms.
637 The top 15 precursors were selected for MS2 in a data dependent manner, within a mass
638 range of m/z 600 – 1600 and a minimum intensity threshold of $5e4$ and an isolation width of
639 1.5 m/z . HCD was performed at 27 % collision energy and detected in the orbitrap with a
640 resolution of 30000 with the first mass at m/z 120, an AGC target of $2e5$ and a maximum
641 injection of 250 ms.

642 EThcD spectra were acquired in a product ion-dependent manner ($[\text{HexNAc}+\text{H}]^+$ -ion) based
643 on the method above. Precursor isolation width was set to 1.2 m/z . Calibrated charge-
644 dependent ETD parameters were used with supplemental activation collision energy of 25%,
645 an AGC target of $2e5$ and a maximum injection time of 350 ms.

646 Glycopeptide identification was performed using Byonic version 3.5 (Protein Metrics Inc.).
647 Trypsin and/or AspN were set as proteases with a maximum of two missed cleavage sites.
648 Mass tolerance for the precursor and HCD fragments was 10 ppm and for EThcD 20 ppm.
649 The glycan database was "O-glycan 78 mammalian". The following modifications were
650 allowed: carbamidomethyl (Cys; fixed), oxidation (Met; variable common 2), pyroglutamine

651 on N-term (Gln, variable, rare 1), acetylation N-term (variable rare1), deamidation (Asn,
652 variable common 2), formylation N-term (variable rare1). Glycopeptides with a score above
653 250 were selected and further manually inspected.

654

655 **scRNA-sequencing**

656 For the expression of ACE2 and receptor genes across different tissues, datasets were
657 retrieved from published datasets in multiple human tissues, including lung and upper
658 airway^{50, 51}, ileum⁵², colon⁵³, pancreas⁵⁴, spleen⁵⁵ and thymus⁵⁶. These datasets are available
659 and can be visualized and assessed through a website portal (www.covid19cellatlas.org)³².
660 The processed .h5ad files were loaded by “read_h5ad” and violin plots were illustrated by
661 “pl.stacked_violin” in scanpy 1.5.1, which is a model for single cell analysis in Python⁵⁷. For
662 the single-cell RNA sequencing analysis of bronchoalveolar immune cells in patients with
663 SARS-CoV-2, dataset was retrieved from Liao et al.¹² and Gene Expression Omnibus (GEO)
664 under the accession number GSE145926, which contains 6 severe and 3 moderate SARS-
665 CoV-2 patients and 3 healthy controls¹².

666 The raw data with .h5 format was loaded for analysis through R package Seurat v3^{58, 59}. For
667 each sample, cells were filtered out if they contain genes less than 200 or more than 6000, or
668 if their UMI is less than 1000, or mitochondrial gene percentage larger than 0.1. The
669 remaining cells were integrated into a gene-barcode matrix and then normalized and log-
670 transformed. We identified 2000 highly variable genes by ‘vst’ method for the downstream
671 PCA analysis. RNA count and the percent of mitochondrial genes were regressed out in the
672 scaling step. We chose the top 50 principal components for the UMAP and graph-based
673 clustering. The cell type identity was referred to Liao et al.¹² and feature genes were
674 demonstrated by “DimPlot”.

675

676 **Data availability**

677 All data is available in the manuscript or in the supplementary information. The single cell
678 RNA-sequencing datasets are available online with the corresponding references.

679

680

681 **Acknowledgments:**

682 This work was supported by National Institutes of Health grants P41GM103694 and R24
683 GM137763, and Evergrande MassCPR subaward 280870.5116795.0025 to E.M., as well as
684 an NIH supported (1UL1TR001430) COVID-19 Related Research Award from the Boston
685 University Clinical and Translational Science Institute (CTSI) to A.E.. SARS-CoV-2 isolate
686 USA_WA1/2020 was kindly provided by CDC's Principal Investigator Natalie Thornburg,
687 nax3@cdc.gov, and the World Reference Center for Emerging Viruses and Arboviruses
688 (WRCEVA). The following reagent was obtained through the NIH AIDS Reagent Program,
689 Division of AIDS, NIAID, NIH: NIH 3T3, NIH 3T3-DC-SIGN+ and NIH 3T3-L-SIGN+
690 from Drs. Thomas D. Martin and Vineet N. KewalRamani. We thank Dr. Jamie Heimburg-
691 Molinaro for help in editing the manuscript.

692

693 **Author contributions:**

694 C.G. and R.D.C. conceived the project. C.G. and Y.M. performed the western blots and
695 ELISA. N.J. and K.S. performed N-glycomics and O-glycoproteomics analysis. J.Z.
696 maintained the cell culture and performed cell binding assays and internalization assay. C.G.
697 performed glycan microarray analysis. H.Z. and J.L. performed scRNA-seq analysis. A.J.H.,
698 E.M., I.D. and J.K. provided important materials for the experiments. K.T., A.E. and R.D.C.
699 supervised experiments. C.G. wrote the first draft of the manuscript and all authors
700 contributed to the final version.

701

702 **Competing interest declaration:**

703 The authors declare no competing interests.

704

705 **Additional information:**

706 Supplementary Information is available for this paper.

707 * Correspondence and requests for materials should be addressed to: Richard D. Cummings,
708 Ph.D., Director, National Center for Functional Glycomics, Department of Surgery, Beth
709 Israel Deaconess Medical Center, Harvard Medical School, CLS 11087 - 3 Blackfan Circle,
710 Boston, MA 02115, Tel: 1-617-735-4643, rcummin1@bidmc.harvard.edu

711

712 **Supplementary Information:**

713 Supplementary Figures 1-9

714 Supplementary Tables 1-3 (in separate excel files)

715

716