The SARS-COV-2 spike protein binds sialic acids, and enables rapid detection in a lateral flow point of care diagnostic device

Alexander N. Baker, a, † Sarah-Jane Richards, a, † Collette S. Guy, a, b Thomas R. Congdon, a Muhammad Hasan, a Alexander J. Zwetsloot, e Anne Straube, f Marc Walker, f Simona Chessa, c Giulia Pergolizzi, c Simone Dedola, c Robert A. Field c, d and Matthew I. Gibson a, e

a) Department of Chemistry, University of Warwick, UK, CV4 7AL; b) School of Life Sciences, University of Warwick, UK, CV4 7AL; c) Iceni Diagnostics Ltd, Norwich Research Park, Norwich, NR4 7GJ; d) Department of Chemistry and Manchester Institute of Biotechnology, University of Manchester, Manchester, UK, M1 7DN; e) Warwick Medical School, University of Warwick, UK, CV4 7AL; f) Department of Physics, University of Warwick, UK, CV4 7AL

ABSTRACT: There is an urgent need to understand the behavior of novel coronavirus (SARS-COV-2), which is the causative agent of COVID-19, and to develop point-of-care diagnostics. Here, a glyconanoparticle platform is used to discover that N-acetyl neuraminic acid has high affinity towards the SARS-COV-2 spike glycoprotein, demonstrating its glycan-binding function. Optimization of the particle size and coating enabled detection of the spike glycoprotein in lateral flow and showed selectivity over the SARS-COV-1 spike protein. Using a viral particle mimic, paper-based lateral flow detection was demonstrated in under 30 minutes showing the potential of this system as a low-cost detection platform.

In December 2019 a novel zoonotic coronavirus (SARS-COV-2), reported in Wuhan (China), led to a pandemic of the respiratory disease COVID-19. There are currently few therapeutics and no vaccine. Diagnostics, surveillance and case isolation are the primary tools for controlling the spread of the virus and driving down the basic reproduction (R0) value. Following the successful genome sequencing of SARS-COV-2, RT-PCR-based (reverse transcription polymerase chain reaction) diagnostics were rapidly developed. These require dedicated laboratory facilities, trained personnel and do not provide an instant output, and while highly specific, false negatives are reported at 3% versus chest CT scans. There are also reports of conflicting RT-PCR results in samples from the same patient. The results can depend on the sampling location i.e. throat versus lower respiratory tract, and the false negative rate is highest during the early stages of infection. An alternative detection platform to RT-PCR is the lateral flow device (LFD), such as the home-pregnancy test, which typically use antibodies as the detection units immobilized to both the stationary phase (e.g. nitrocellulose paper) and the mobile phase (e.g. gold nanoparticle), forming a ‘sandwich’ with the analyte. Test lines show a positive (red line) response by eye. Such LFDS require little or no clinical infrastructure or training and they can be used in the patient’s home. The cost-effectiveness of these inexpensive devices has been demonstrated by various studies of malaria rapid diagnostic tests and were found to compare well to the more expensive RT-PCR for Ebola diagnosis. In addition to antibodies, other biological recognition units such as nucleic acids, glycans and lectins could be used. Glycans have not been widely applied in lateral flow, but offer new opportunities and advantages compared with antibody-based systems. Glycans have reduced cold chain needs compared to proteins and are thus suited to low resource, triage or emergency settings. A further benefit of glycans would be the detection of intact virus. For SARS-COV-2, viral RNA (e.g. from a positive RT-PCR result) is detected past the point where patients are no longer infectious, resulting in extended hospital stays. Glycans (carbohydrates) direct myriad binding and recognition events in biology from cell-cell communication to being markers of disease. Analysis of the 2009 swine (zoonotic) influenza pandemic showed that porcine viral hemagglutinins, which normally bind α2,3’-linked sialic acids, switched to binding α2,6’-linked sialic acids found in human respiratory tracts. This demonstrates the importance of glycan-binding during infection. This switch in glycan specificity has enabled the establishment of glycan-functional biosensors to rapidly identify strains without the need for genetic approaches. All coronaviruses display homotrimers of spike glycoproteins on their surface. Sialic acid binding by the S1 spike protein subunits is crucial for coronavirus to engage host cells, whilst the S2 domain initiates viral fusion. Tortorici et al showed the structural basis for 9-O-acetylated sialic acid binding to a human coronavirus (strain OC43) by cryo-EM, and affinity to this ligand to the HKU1-HE strain, has also been reported. MERS S1 preferentially binds α2,3’-linked over α2,6’-linked sialic acids but any acetylation of the sialic acids decreases affinity, which is distinct from OC43. This evidence shows that sialic acid binding is crucial in coronavirus infection and potentially in its zoonosis to human hosts, but that the exact glycan partner can vary between strains. The glycobiology of coronaviruses has not yet been explored in detail, but the above examples demonstrate that glycan ‘anchoring’ of coronaviruses may offer opportunities for detection using capture techniques such as LFD. Individual glycans display low affinity to their protein targets, but this is overcome in nature by multiple copies of glycans found in intact virus. Using multivalent gold nanoparticles bearing sialic acid derivatives and their interaction with the spike glycoprotein from SARS-COV-2. We find that α-N-acetyl neuraminic acid binds the spike glycoprotein strongly, and subsequently exploit this interaction as the detection unit in a prototype lateral flow diagnostic, which requires no centralized infrastructure. Figure 1A shows the sequence alignments of coronavirus spike proteins from SARS-COV-1, IBV, MERS-COV, and SARS-COV-2 with respect to the known sialic acid binding groove of HCOV-OC43. There are no clearly conserved residues between the sequences but Phel91 and Pro94 are common to all, apart from MERS. This is in marked contrast
to the spike S protein, which is often highly conserved.\textsuperscript{30} This lack of sequence homology within the sialic acid binding groove may contribute to the ability of the virus to cross between species.\textsuperscript{31} The predicted binding groove of SARS-COV-2 is shown in Figure 1B. Figure 1C shows the MERS sialic acid binding site in complex with α2,3\textsuperscript{*} sialyllactose. Crucially, only the neuraminic acid unit, not the lactose, is engaged. This is in contrast to influenza hemagglutinins which display contacts to the galactose residues.\textsuperscript{32} This evidence suggests that N-acetyl neuraminic acid is a reasonable target for SARS-COV-2 binding and hence a potential capture ligand for a new ‘glyco-LFD’ diagnostic device.

Figure 2A shows a design schematic for a glyco-LFD. Typical LFDs use antibodies, but here the glycan is immobilized on the test strip and also in the mobile phase onboard gold nanoparticles, providing multivalency (and hence affinity), for dissecting SARS-COV-2 binding and for the LFD. Our nanoparticle design concept uses telechelic polymer tethers which conjugate the glycans, by displacement of an α-terminal pentafluorophenyl (PFP) group, and immobilization onto gold particles via the α-terminal thiol. Poly(α-hydroxyethyl acrylamide), PHEA, was chosen as the polymer to give colloidally stable particles and as an acrylamide, it is not easily hydrolysable unlike acrylates for example, Figure 2B. Table 1.\textsuperscript{32,33} PHEA was synthesized using RAFT (reversible addition-fragmentation chain transfer) polymerization resulting in dispersities below 1.3. The PHEA lengths were selected on the basis of performance in initial lateral flow screening assays (data not shown) and from reports of their colloidal stability.\textsuperscript{32,33} Amino-glycans were synthesized by reduction of anomeric azides and subsequently conjugated to the PHEAs by displacement of the PFP group, which was confirmed by \textsuperscript{19}F NMR. Polymers were then assembled onto citrate-stabilized gold nanoparticles and excess ligand removed by centrifugation/resuspension cycles. The nanoparticles were characterized by UV-Vis, dynamic light scattering (DLS), transmission electron microscopy (TEM) and XPS (x-ray photoelectron spectroscopy) to confirm surface coating (Table 2 and S1). Following observations via DLS and UV-Vis that 16 nm sialyllactose particles were less stable than 35 nm particles, the latter were selected for initial glycan-binding assays.

![Figure 1](image1.png)

**Figure 1.** A) Sequence alignment of hypothesized sialic acid binding sites of coronaviruses; B) Model showing the hypothesized sialic acid binding site for SARS-COV-2 spike protein assembly and the S1, S2 domains; C) MERS-Sialic acid binding site in complex with α2,3\textsuperscript{*} sialyllactose (from PDB entry 6Q06\textsuperscript{34}).

![Figure 2](image2.png)

**Figure 2.** Design concept for glyco-lateral flow devices. A) Lateral flow assay for virus, using glycan capture units; B) Synthetic procedure for glyconanoparticles.

### Table 1. Polymer Characterization

<table>
<thead>
<tr>
<th>Code</th>
<th>M:CTA</th>
<th>M\textsubscript{SEC}(a) (g.mol\textsuperscript{-1})</th>
<th>M\textsubscript{NMR}(b) (g.mol\textsuperscript{-1})</th>
<th>D(c) (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHEA\textsubscript{58}</td>
<td>20</td>
<td>2800</td>
<td>5100</td>
<td>5000</td>
</tr>
<tr>
<td>PHEA\textsubscript{58}</td>
<td>25</td>
<td>3400</td>
<td>6400</td>
<td>5500</td>
</tr>
<tr>
<td>PHEA\textsubscript{58}</td>
<td>30</td>
<td>4000</td>
<td>7200</td>
<td>6700</td>
</tr>
</tbody>
</table>

(a) Estimated from [M]:[CTA]; (b) From DMF SEC verses PMMA standards; (c) \textsuperscript{1}H NMR end-group analysis.

### Table 2. Nanoparticle Characterization

<table>
<thead>
<tr>
<th>Code</th>
<th>U\textsubscript{Max}(a) (nm)</th>
<th>A\textsubscript{SPR}/A\textsubscript{Vis}(b)</th>
<th>D\textsubscript{H2O}(c) (nm)</th>
<th>D\textsubscript{DLS}(d) (nm)</th>
<th>D\textsubscript{TEM}(e) (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AuNP\textsubscript{16}</td>
<td>519</td>
<td>1.64</td>
<td>16</td>
<td>20.7 ± 0.8</td>
<td>14 ± 2</td>
</tr>
<tr>
<td>NeuNAC-PHEA\textsubscript{40}AuNP\textsubscript{16}</td>
<td>527</td>
<td>1.66</td>
<td>16</td>
<td>40.9 ± 0.6</td>
<td>(-)</td>
</tr>
<tr>
<td>NeuNAC-PHEA\textsubscript{40}AuNP\textsubscript{16}</td>
<td>526</td>
<td>1.68</td>
<td>18</td>
<td>44.2 ± 0.8</td>
<td>(-)</td>
</tr>
<tr>
<td>AuNP\textsubscript{35}</td>
<td>526</td>
<td>1.91</td>
<td>35</td>
<td>34.5 ± 0.5</td>
<td>35 ± 3</td>
</tr>
<tr>
<td>NeuNAC-PHEA\textsubscript{40}AuNP\textsubscript{35}</td>
<td>531</td>
<td>1.98</td>
<td>45</td>
<td>46.2 ± 0.7</td>
<td>(-)</td>
</tr>
<tr>
<td>NeuNAC-PHEA\textsubscript{40}AuNP\textsubscript{35}</td>
<td>531</td>
<td>1.99</td>
<td>45</td>
<td>55.3 ± 0.8</td>
<td>(-)</td>
</tr>
</tbody>
</table>

(a) SPR absorption maximum; (b) Absorbance ratio of SPR to 450 nm; (c) Estimated from UV-Vis;\textsuperscript{38} (d) From dynamic light scattering; (e) From TEM, from average of >100 particles, showing ±S.D.

With the glyconanoparticles in hand, recombinant S1 subunit (SARS-COV-2,S1) spike protein was immobilized onto biolayer interferome-
try (BLI) sensors\textsuperscript{18} to replicate a lateral flow situation, which is the primary aim of this work. Since there are 22 N-linked glycans per protein which are not present in bacteria-expressed protein,\textsuperscript{39} we used protein expressed in mammalian cells (HEK) to ensure glycosylation. Figure 3 A shows \(\alpha\),NeuNAc-AuNPs bind to a greater extent compared to both sialylactose isomers (\(\alpha2,3\),\(\alpha2,6\)) and the monosaccharide control (glucose). X-ray photoelectron spectroscopy analysis (SI) of these particles revealed that the NeuNAc/Glc monosaccharide-terminated polymers had a higher grafting density than the sialylactose trisaccharides by a ratio of 2:1 (35 nm)/ 3:1 (16 nm), due to the difference in glycan size. It is therefore important to note that this data does not rule out sialylactose binding (and indeed, in LFD we do see binding), but that in this system NeuNAc gave the strongest response and consequently was taken forward. The observation of strong binding of NeuNAc agrees with the MERS spike protein structure in complex with sialylactose, where only the NeuNAc, not the galactose linker has significant contacts. Whilst outside the scope of this work, the identification of the sialic acid binding function of the spike protein may provide fundamental guidance as to how the virus engages host cells, or is processed within them, and this (to the best of our knowledge) is the first report of this matter.

The next step was to evaluate the impact of particle size on binding. Both 16 and 35 nm gold (relevant diameters for LFDs) NeuNAc particles were used to interrogate SARS-COV-2,SI, Figures 3B/C. End point dose dependency (Fig 3D) showed similar binding trends for both particles, with an apparent \(K_a \approx 1\) nM, noting that for multivalent systems exact \(K_a\)'s cannot be extracted. The plots are shown in terms of OD (SPR absorption maximum) as this is standard for AuNP concentration.

With the successful identification of NeuNAc as a target ligand, its application as the capture unit in lateral flow was examined. The performance of an LFD depends upon not only the affinity of the capture ligand (NeuNAc) but on the flow of the particles. ‘Half’ lateral flow assays (Figure 4A) were established to optimize the particles. The test line was either 2,3’-sialyllactose-BSA (negative control for non-specific binding) or immobilized SARS-COV-2,SI, and the nanoparticles were flowed over them (all original lateral flow strips are in SI). Please note all figures have representative data for each condition, but additional data is available in SI. Importantly, all particles bound SARS-COV-2,SI showing this detection method is valid. 16 nm Particles and NeuNAcPHEA\textsubscript{0}@AuNPs gave stronger signals than the other 35 nm particles, but also more background, hence NeuNAcPHEA\textsubscript{0}@AuNPs\textsubscript{35} particles were used from this point onwards, Figure 4B/C. \(\alpha2,3\)‘ and \(\alpha2,6\)’sialyllactosamine particles were also tested, and on larger AuNPs (55-70 nm) and longer polymers (PHEA\textsubscript{2}) too, but gave no improvement over NeuNAc-PHEA\textsubscript{0}@AuNPs\textsubscript{35}. Blocking of the NeuNAc particles with BSA before running was also explored in an attempt to further reduce the background, as is common in LFDs. BSA blocking did not improve the performance of the NeuNAc systems but it did reduce off-target binding in the \(\alpha2,3\)’ sialyllactosamine systems tested. Encouraged by these results, the specificity and function of the NeuNAcPHEA\textsubscript{0}@AuNPs\textsubscript{35} particles were tested against a panel of test-line immobilized lectins [1 mg.mL\textsuperscript{-1}]. Total signal intensity is plotted in Figure 4D confirming that NeuNAc-AuNPs have no non-specific binding. The only lectin which bound was RCA\textsubscript{20}, which is known to have some affinity towards sialic acids.\textsuperscript{40} To test binding specificity in a more challenging scenario, the particles were screened against the spike protein, SARS-COV-2,SI, (the desired target) and also against the S1 spike domain of a previous zoonotic coronavirus SARS-COV-1,\textsuperscript{23} responsible for 2003 ‘SARS’ outbreak. As can be seen in Figure 4E, the NeuNAc particle system has a clear preference for SARS-COV-2 demonstrating selectivity in this glyconanoparticle system. While this does not rule out binding, it does show the particles/glycan do not generate sufficient signal against SARS-COV-1. This data further supports the notion that the terminal NeuNAc is the key binding motif.

To explore the detection limits and specificity of this system, NeuNAc (positive) and galactose (negative) nanoparticles were screened against a dilution series of SARS-COV-2,SI (Figure 5A/B). At the highest concentration (0.5 mg.mL\textsuperscript{-1}) galactose particles showed weak binding to SARS-COV-2,SI. NeuNAc particles showed significantly stronger binding, with an apparent limit of detection below 8 µg.mL\textsuperscript{-1}. Encouraged by successful binding, a complete dipstick sandwich assay was established where the analyte was added to the gold particle solution, rather than dried onto the nitrocellulose paper. The test line was NeuNAc-BSA (validated to capture the particles by BLI, SI) and RCA\textsubscript{20} as a control line, which is essential in lateral flow devices to ensure each device is functional. To mimic the virus, spike glycoprotein was immobilized onto 100 nm polystyrene nanoparticles which
match the diameter of the coronavirus. Figure 5C shows the results of testing this system (original lateral flow strips with no image enhancement are in the SI). The lateral flow devices could clearly detect the virus-like particles at a concentration of just 5 µg.mL⁻¹ (5 nM) protein, which is in line with the detection limits from Figure 5B. Controls using naked polystyrene colloids showed no binding to the test line, ruling out non-specific interactions, and a control (with no polystyrene analyte) only showed control line binding. The resolution of the test spots could be further enhanced using a simple silver-staining protocol⁴⁴ which improves the ‘by eye’ detection, Figure 5C. An additional control of two influenza strains (which bind sialyllactoses) were shown to have little off-target binding as influenza hemagglutinins require the galactose linker in addition to the sialic acid, for strong binding (SI).⁴⁴ This data strongly supports the hypotheses that targeting the sialic acid binding site in SARS-COV-2 using glycosylated nanoparticles is a valid route for rapid point of care diagnosis of this emergent virus. Future work will be required to validate viral load detection limits, and optimize/formulate the LFD device to increase signal, but the principle is clearly proved in this work.

In conclusion, we have demonstrated an entirely glycan-based lateral flow detection system that can be used to recognize the spike glycoprotein from the SARS-COV-2 virus in under 30 mins. Guided by sequence alignment against other coronavirus spike proteins it was hypothesized that sialic acids may bind this protein, to enable capture/detection. Using a nanoparticle based biolayer interferometry platform we demonstrated that α,N-acetyl neuraminic acid is a ligand for the spike glycoprotein. The gold nanoparticles and polymer tethers (for glycan capture) were optimized and it was found in ‘half sandwich’ lateral flow assays that α,N-acetyl neuraminic acid particles have high selectivity towards the SARS-COV-2 spike protein, including specificity over SARS-COV-1 and a panel of lectins. Guided by this, we successfully detected a virus mimic particle bearing SARS-COV-2.S1 in under 30 mins, with a detection limit of the spike protein around 5 µg.mL⁻¹. This work provides proof that glycan binding can be exploited to create rapid point-of-care diagnostics in a format which requires no infrastructure and limited training and, to the best of our knowledge, is the first reported all-glycan lateral flow system. This approach may find application for disease surveillance or mass testing at transport/work hubs or even for self/home testing. Finally, the observation that SARS-COV-2 can engage sialic acids found on human respiratory cells may provide insight into its zoonosis and infection pathways to help guide new interventions.

ASSOCIATED CONTENT
Supporting Information
The Supporting Information (SI) is included at the bottom of this file. This includes experimental details, characterization, BLI curves, unenhanced lateral flow device photos and additional control experiments.

AUTHOR INFORMATION
Corresponding Author
* m.i.gibson@warwick.ac.uk

Author Contributions
The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript. ‡ These authors contributed equally.

Conflict of Interest Statement
RF is a shareholder and CEO of Iceni Diagnostics Ltd. ANB, SJR, MIG are named inventors of a patent application relating to this work.

ACKNOWLEDGMENT
MIG is supported by the ERC (638661) and Royal Society (Industry Fellowship 191037). ANB thanks the BBSRC MIBTP program (BB/M01116X/1) and Iceni Diagnostics Ltd. UoW, EPSRC (EP/R511808/1) and BBSRC (BB/S06783/1) impact acceleration accounts are thanked. AJZ is funded by the MRC DTP (MR/N014294/1). AS is a Wellcome Trust Investigator (200870/Z/16/Z). The Warwick Polymer and Electron Microscopy Research Technology Platforms (Y. Han) are acknowledged for the SEC/EM analysis. Dr C. Biggs is thanked for lateral flow device preparation. We sincerely thank the technical and administrative staff of the UoW who enabled our laboratory to remain open during the COVID-19 pandemic. Prof. Peter Scott is thanked for critically reading this manuscript.

REFERENCES


