Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Protein Expression and Purification 183 (2021) 105861

Contents lists available at ScienceDirect

Protein Expression and Purification

journal homepage: www.elsevier.com/locate/yprep

Expression of SARS-CoV-2 surface glycoprotein fragment 319–640 in *E. coli*, and its refolding and purification

Gabriel A. Fitzgerald [a], Andrei Komarov [a], Anna Kaznadzey [a,b], Ilya Mazo [a,c], Maria L. Kireeva [a,*]

[a] VirIntel, LLC, Frederick, MD, USA
[b] A. A. Kharkevich Institute for Information Transmission Problems, Russian Academy of Sciences, Moscow, Russia
[c] Argentys Informatics, LLC, Gaithersburg, MD, USA

ARTICLE INFO

Keywords:
COVID-19
Spike
Antibody test
Antigen
Bacterial expression
Receptor-binding domain

ABSTRACT

Sensitive and specific serology tests are essential for epidemiological and public health studies of COVID-19 and for vaccine efficacy testing. The presence of antibodies to SARS-CoV-2 surface glycoprotein (Spike) and, specifically, its receptor-binding domain (RBD) correlates with inhibition of SARS-CoV-2 binding to the cellular receptor and viral entry into the cells. Serology tests that detect antibodies targeting RBD have high potential to predict COVID-19 immunity and to accurately determine the extent of the vaccine-induced immune response. Cost-effective methods of expression and purification of Spike and its fragments that preserve antigenic properties are essential for development of such tests. Here we describe a method of production of His6-tagged S319-640 fragment containing RBD in *E. coli*. It includes expression of the fragment, solubilization of inclusion bodies, and on-the-column refolding. The antigenic properties of the resulting product are similar, but not identical to the RBD-containing fragment expressed in human cells.

CRediT statement

Gabriel A. Fitzgerald: Investigation, Writing – original draft. Writing – review & editing. Andrei Komarov: Investigation. Anna Kaznadzey: Data curation. analysis, Writing – review & editing. Ilya Mazo: Conceptualization, Project administration, Funding acquisition; Maria L. Kireeva: Conceptualization, Supervision, Investigation, Writing – original draft, Writing – review & editing

1. Introduction

Efficient public health response to the developing COVID-19 pandemic requires expanding diagnostic tools for revealing seroconversion (antibody development and release into the blood stream) to SARS-CoV-2 infection. Sensitive and specific antibody tests are essential for epidemiological and public health studies. They can be used to establish the extent of an outbreak, map its overall geographical distribution and the hotspot locations, and identify groups that are at higher risk of infection. This data serves as a foundation for public health measures and control strategies [1]. Vaccine research and development is another important area of serology test application [2,3].

Antigen selection is crucial for the development of an effective serology test. Recent observations suggest that the presence of antibodies targeting SARS-CoV-2 surface glycoprotein (S or Spike) and its receptor-binding domain (RBD) better correlate with virus neutralization *in vitro* than the presence of antibodies to other SARS-CoV-2 proteins [4]. The relative abundance of circulating Spike-specific immunoglobulins positively correlates with the survival and recovery of COVID-19 patients [5]. All COVID-19 vaccine candidates are based on Spike or its fragments as antigens [6]. Therefore, serological tests utilizing Spike or its fragments as antigens have a higher potential to predict COVID-19 immunity resulting from vaccination or prior infection.

VirIntel COVID-19 Antibody test is based on two antigens to enhance the antibody detection [7]. A nucleocapsid protein, an abundant and highly antigenic component of SARS-CoV-2 virus particle, is one of the antigens used in the test. A fragment of Spike encoding RBD followed by 50 extra amino acids (S319-591) forming the Spike subdomain SD1 [8,9] is used as the second antigen. In the clinical validation study, all 31 blood serum samples collected from individuals recovered from COVID-19 showed the presence of IgGs specific to S319-591. 88 out of 89 blood serum samples collected before the COVID-19 pandemic started did not contain antibodies binding S319-591.

To further improve the sensitivity and specificity of the test, we

* Corresponding author.
E-mail address: maria.kireeva@virintel.com (M.L. Kireeva).

https://doi.org/10.1016/j.pep.2021.105861
Received 28 September 2020; Received in revised form 19 February 2021; Accepted 24 February 2021
Available online 2 March 2021

explored the possibility to extend the Spike fragment used as the antigen. Amino acid residues S592-640 form a part of subdomain SD2, which, according to our bioinformatics analysis, might contain an additional B-cell epitope. This region also contains an amino acid residue D614, mutation in which emerged and rapidly spread early in the pandemic [10]. We aimed to examine the antigenic properties of the surface glycoprotein fragment S319-640. S319-591 fragment was efficiently expressed in transiently transfected HEK293 cells, similar to RBD-containing fragments used in other serology tests [11–13]. However, S319-640 fragment was not detected in the conditioned media of HEK293 cells transiently transfected with the corresponding expression construct. Therefore, we attempted expression of this SARS-CoV-2 Spike fragment in *E. coli*.

2. Materials and methods

2.1. Protein expression and extraction

BL21 (DE3) *E. coli* cells were transformed with a pET30a + expression vector encoding residues 319–640 of SARS-CoV-2 surface glycoprotein. The cloning strategy of the codon-optimized sequence is shown in Supplemental Fig. S1. Cells were grown to an OD_{600} of 0.5 and induced with 0.2 mM IPTG at 37°C. After 2 h the cells from 1 L of culture were harvested at 10,000 rpm in a Beckman JA-10 rotor for 10 min and solubilized with 8 mL BugBuster (Millipore Sigma) per liter of culture for 30 min at 30°C. 15 mL Base Buffer (20 mM Tris pH 8, 0.5 M NaCl, 10% glycerol and 5 mM β-mercaptoethanol) per liter of culture was added to the solubilized cells and sonicated on ice 15 × 20 s with a 50% duty cycle at 75% power. The resulting suspension was centrifuged for 10 min at 10,000 rpm at 4 °C. The pellet was resuspended in 15 ml per liter of culture in Base Buffer supplemented with 2 M urea and homogenized with a glass dounce homogenizer. The sonication and centrifugation procedures were repeated, and the resulting pellet was resuspended by dounce in 15 mL per liter of culture of Base Buffer supplemented with 6 M urea. The pellet was again subjected to the same sonication and centrifugation procedures and the supernatant was collected.

2.2. Protein refolding

The extracted supernatant was diluted 10-fold in Refolding Buffer (20 mM Tris pH 8, 20% glycerol, 55 mM glucose, 0.5 M NaCl, 2 mM reduced glutathione (GSH) and 0.2 mM oxidized glutathione (GSSG)) supplemented with 2 M urea. The diluted extract was applied to a HisTrap HP column (Millipore Sigma) at 1 mL/min at room temperature. The column was washed with 2 column volumes (CV) of Refolding Buffer supplemented with 1.5 M urea at 0.2 mL/min for the first 0.5 CV and 0.6 mL/min for the remaining 1.5 CV. This procedure was repeated with Refolding Buffer supplemented with 1 M urea, 0.5 M urea and 0 M urea. Following the refolding procedure, the column was washed with Refolding Buffer supplemented with 30 mM imidazole and eluted with the same buffer supplemented with 0.5 M imidazole.

2.3. Enzyme-linked immunosorbent assays

ELISA setup was based on the procedure described by Amanat et al. [13] The 96-well Immulon 4 HBX plates (ThermoFisher Scientific #3855) were coated with 0.1 μg of protein dissolved in phosphate buffered saline (PBS), pH 7.4 per well overnight. The plates were blocked with 3% non-fat milk in PBS, pH 7.4, containing 0.1% Tween-20 (PBS-T), washed with PBS-T, and incubated with serum samples diluted 1:50 or 1:150 in PBS-T containing 1% dry milk. A human chimeric SARS-CoV-2 spike S1 monoclonal antibody HC2001 (Genscript #A02038) at 0.1 μg/mL was used as a positive control. The bound IgG antibodies were detected using horseradish peroxidase (HRP)-conjugated monoclonal antibodies recognizing an Fc domain of human IgG (GenScript, # A01854).

The blood serum samples were collected from donors recovered from COVID-19 (the diagnosis confirmed by an RT-PCR test) and their family members. The blood draws were performed after the symptoms of the disease subsided and over 4 weeks after the disease onset.

3. Results and discussion

The SARS-CoV-2 surface glycoprotein fragment (S319-640) was expressed at high levels in BL21 (DE3) cells (Fig. 1a) but was found to be primarily contained within inclusion bodies as indicated by the firm white pellets obtained after cell lysis (data not shown). Extraction of the protein from inclusion bodies was not effective with urea alone (Fig. 1a) but was effective with high concentrations (6 M) of guanidium chloride (Fig. 1b). Despite this effective solubilization, the protein could not be recovered upon removal of guanidium chloride, even with intermediate concentrations of urea to aid in refolding (Fig. 1b).

BugBuster is a detergent based solubilization reagent that has been shown to be effective in extraction of proteins from inclusion bodies [14]. Incubation with BugBuster, in combination with sonication steps in the presence of increasing concentrations of urea allowed for extraction of S319-640 from *E. coli* (Fig. 1c). The protein extracted from this procedure was stable for several weeks at 4°C in 6 M urea. However, decreasing the urea concentration below 2 M resulted in protein precipitation, thus calling for development of an optimized refolding procedure.

The protocol for the refolding of S319-640, adapted from Zhao et al. [15] utilized gradual decreases in urea concentrations in combination with oxidized and reduced glutathione to allow for proper formation of disulfide bonds along with glucose and glycerol as stabilizing agents to refold a fragment of the SARS-CoV-1 surface glycoprotein. We devised a Ni^{2+}-NTA column-based refolding procedure to allow for simultaneous purification and refolding of the S319-640 protein. In this procedure the protein solution was applied to the column in the presence of 2 M urea, after which the urea concentration was stepped down in 0.5 M increments with slow flow rates (0.2–0.6 mL/min) until there was no remaining urea.

A significant amount (>50%) of the protein was contained in the flow-through (Fig. 1d, lane 2), likely due to the presence of contaminating inclusion bodies that were not fully solubilized. The column was washed with 30 mM imidazole (lane 3) following the refolding procedure and eluted with 0.5 CV fractions of 0.5 M imidazole in refolding buffer (lanes 4, 5 and 6 of Fig. 1d). About 10% of the loaded protein was eluted from the column in non-denaturing conditions. The protein obtained from refolding procedure remained soluble at 4°C for several months and was successfully used for detection of SARS-CoV-2 antibodies in human blood serum samples (Fig. 2).

The refolded recombinant surface glycoprotein fragment S319-640 interacts with the antibodies specific to SARS CoV-2 in a manner distinct from that of the S319-591 fragment expressed in human cells. Fig. 2 summarizes results of two independent experiments, in which several human blood serum samples were analyzed.

In Fig. 2a, a comparison of the signals from a no-serum-added negative control and S1-specific human chimeric monoclonal antibody (mAb) HC2001 positive control was used to demonstrate the dynamic range of the assay. On the hRBD-coated plates (blue bars), the difference between the signals in the negative and positive controls was about 12-fold. However, bRBD (orange bars) barely bound the monoclonal antibody: the signal from the presumptive positive control with bRBD is less than two-fold higher than the signal in the negative control. It is comparable to the signals from samples 13579 and 7489, which were collected from asymptomatic individuals, and have never shown any reactivity to either SARS-CoV-2 nucleocapsid protein or to S319-591 surface glycoprotein fragment in any of our experiments [7]. This observation suggests that bRBD is lacking the epitope recognized by HC2001 monoclonal antibody used in this study.

Nevertheless, testing of human blood serum samples reveals that

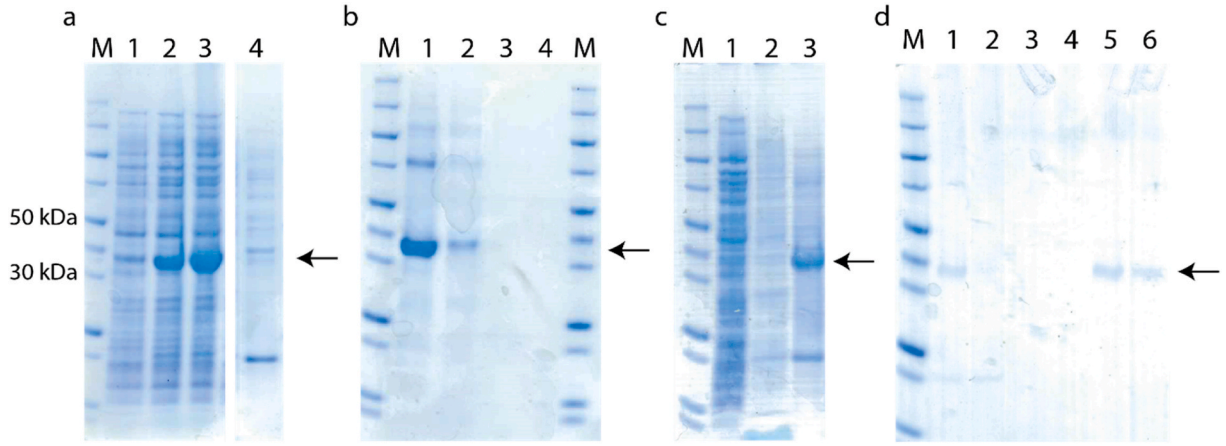

Fig. 1. Expression, purification and refolding of S319–640. a) Expression of S319-640 (arrow) in BL21 (DE3) in the absence of IPTG (lane 1) or in the presence of 0.2 mM IPTG for 1 h (lane 2) or 2 h (lane 3) with no protein observed after extraction in 6 M urea in the absence of BugBuster reagent (lane 4). **b**) S319-640 (arrow) extracted in 6 M guanidium chloride diluted in 6 M urea (lane 1) subsequently applied to a Na^{2+}-NTA column. Protein was primarily contained in the flow-through (lane 2) with none observed in the column wash in 6 M urea (lane 3) or elution with 0.2 M imidazole (lane 4). **c**) Solubilization of S319-640 (arrow) with BugBuster reagent in combination with successive urea washes and sonication with the whole cell lysate supernatant (lane 1) and 2 M urea wash supernatant (lane 2) lacking in protein but observation of successful extraction in the 6 M urea wash supernatant (lane 3). **d**) On-column refolding of S316-640 (arrow) first diluted to 2 M urea (lane 1) and applied to a His-Trap affinity purification column.

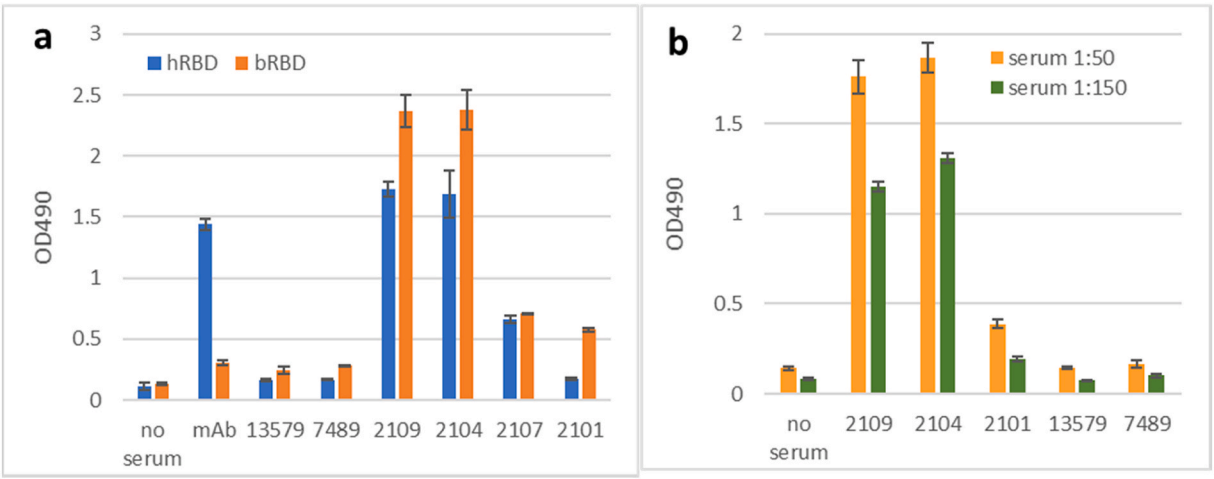

Fig. 2. Antigenic properties of the refolded surface glycoprotein fragment. 4- and 5-digit numbers below the chart correspond to the specimen identification numbers. mAb indicates a positive control with human chimeric SARS-CoV-2 Spike S1 antibody. The data represent average of three technical repeats for bRBD and two repeats for hRBD, and error bars indicate standard deviation. **a,** comparison of IgG binding to the refolded 319–640 SARS-CoV-2 surface glycoprotein fragment (bRBD) and to 319–591 fragment expressed in human cells (hRBD). Serum was diluted 1:50. **b**, effect of serum dilution on the amounts of antibodies bound to bRBD.

bRBD has strong antigenic properties. Samples 2109 and 2104 were collected from a male and a female, a couple who recovered from COVID-19. For both, the diagnosis was confirmed by a PCR test. Both demonstrated strong signals for antibodies binding to RBD fragments expressed in either human or bacterial cells (Fig. 2a) and to the nucleoprotein [7].

The asymptomatic young children, a male (2107) and a female (2101), despite being SARS-CoV-2 negative according to a PCR test performed at the same time as their parents' PCR test, have shown the presence of antibodies to SARS-CoV-2. 2107 has shown the presence of antibodies specific to S319-591 expressed in human cells and to S319-640 expressed in *E. coli*. 2101 has shown presence of antibodies specific only to the latter. The signals from 2107 to 2101 are noticeably lower than those from 2104 to 2109, in agreement with a well-established correlation of the strength of the immune response with the severity of the disease [3]. Interestingly, 2101 serum does not have any reactivity to S319-591 expressed in HEK293 cells (Komarov et al., 2021 [7] and Fig. 2a), but clearly shows reactivity to S319-640 expressed in *E. coli*. To confirm the specificity of this apparent binding, we tested if it remains higher than the signal from the control 13579 and 7489 serum samples when serum is additionally diluted. Fig. 2b compares IgG binding to bRBD at serum dilutions 1:50 (orange columns) and 1:150 (green columns). At 1:150 serum dilution, the signal from 2101 serum specimen still exceeds the signals from a no-serum control and non-reactive serum samples, suggesting that 2101 sample indeed contains IgGs specifically interacting with bRBD.

In summary, we observed that two SARS-CoV-2 Spike RBD fragments, S319-591 (hRBD), expressed in human cells, and S319-640 (bRBD), expressed in *E. coli* in insoluble form and refolded during purification, both specifically bind IgGs induced by SARS-CoV-2 infection. The spectrums of antibodies recognized by these two antigens are similar, but not identical. The similarity is evident from the absence of IgG binding from the COVID-19-negative serum specimens and high amounts of bound IgGs from the COVID-19-positive specimens. One striking difference between bRBD and hRBD is that mAb HC2001 interacts with hRBD, but not with bRBD. The absence of binding could

have been caused by improper folding of bRBD or by absence of glycosylation of the recombinant protein fragment expressed in *E. coli*. The latter appears more likely, because bRBD efficiently binds IgG from COVID-19-positive serum samples. Furthermore, we identified a serum specimen that has a higher reactivity to bRBD, compared to hRBD. A plausible explanation of this observation could be that 2101 serum specimen contains antibodies specific to S592-640 region lacking in hRBD.

The procedure for purification of SARS-CoV-2 Spike fragment 319–640 from inclusion bodies in *E. coli* described here can be applied to any Spike fragment. Recently, S318-510 fragment was expressed in *E. coli* and successfully used in ELISA assays [16]. Our work suggests that this technique can be applied to longer Spike fragments, including those that are difficult to express in mammalian cells. Beside its cost-effectiveness, he main advantage of *E. coli* as an expression host compared to eukaryotic cells (mammalian and yeast alike) is the short time required for transformation, cell propagation, and induction of protein expression. Rapid methods of antigen isolation gain more and more significance as new variants of SARS-CoV-2 carrying mutations in Spike emerge in the population [17]. While D614D mutation does not appear to alter antigenic properties of Spike [18], the effects of other mutations require further investigation. Identification of Q677, an amino acid residue undergoing mutations in multiple emerging SARS-CoV-2 variants [19] suggests that extensive analysis of antigenic properties of the RBD-containing fragments extended beyond S640 amino acid residue of Spike, and fragments of the S592-685 region lacking RBD, might be called for in the nearest future. *E. coli* expression system appears most suitable for the antigen screening purposes, providing an inexpensive and, most importantly, fast way to express, isolate, and analyze multiple protein fragments.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.pep.2021.105861.

References

[1] Peeling, R. W. et al. Serology testing in the COVID-19 pandemic response. Lancet Infect. Dis. 20, e245-e249, doi:10.1016/S1473-3099(20)30517-X.(2020). PMC7367660.

[2] Nie, J. et al. Establishment and validation of a pseudovirus neutralization assay for SARS-CoV-2. Emerg. Microb. Infect. 9, 680-686, doi:10.1080/22221751.2020.1743767.(2020).PMC7144318.

[3] Okba, N. M. A. et al. Severe acute respiratory syndrome coronavirus 2-specific antibody responses in coronavirus disease patients. Emerg. Infect. Dis. 26, 1478-1488, doi:10.3201/eid2607.200841.(2020).PMC7323511.

[4] K.M. McAndrews, et al., Heterogeneous antibodies against SARS-CoV-2 spike receptor binding domain and nucleocapsid with implications for COVID-19 immunity, JCI Insight 5 (2020), https://doi.org/10.1172/jci.insight.142386.

[5] Atyeo, C. et al. Distinct early serological signatures track with SARS-CoV-2 survival. Immunity 53, 524-532 e524, doi:10.1016/j.immuni.2020.07.020.(2020). PMC7392190.

[6] Amanat, F. & Krammer, F. SARS-CoV-2 vaccines: status report. Immunity **52**, 583-589, doi:10.1016/j.immuni.2020.03.007.(2020).PMC7136867.

[7] Komarov, A., Kaznadzey, A., Li, Y., Kireeva, M. & Mazo, I. Dual-antigen system Allows elimination of false positive results in COVID-19 serological testing. Diagnostics 11, doi:10.3390/diagnostics11010102.(2021).PMC7827312.

[8] P. Adhikari, et al., Intra- and intermolecular atomic-scale interactions in the receptor binding domain of SARS-CoV-2 spike protein: implication for ACE2 receptor binding, Phys. Chem. Chem. Phys. (2020), https://doi.org/10.1039/d0cp03145c.

[9] L. Premkumar, et al., The receptor binding domain of the viral spike protein is an immunodominant and highly specific target of antibodies in SARS-CoV-2 patients, Sci Immunol 5 (2020) PMC7292505, https://doi.org/10.1126/sciimmunol.abc8413.

[10] Korber, B. et al. Tracking changes in SARS-CoV-2 spike: evidence that D614G increases infectivity of the COVID-19 virus. Cell 182, 812-827 e819, doi:10.1016/j.cell.2020.06.043.(2020).PMC7332439.

[11] Mehalko, J. et al. Improved production of SARS-CoV-2 spike receptor-binding domain (RBD) for serology assays. Protein Expr. Purif. 179, 105802, doi:10.1016/j.pep.2020.105802.(2021).PMC7687410.

[12] O. Farnos, et al., Rapid high-yield production of functional SARS-CoV-2 receptor binding domain by viral and non-viral transient expression for pre-clinical evaluation, Vaccines 8 (2020) PMC7712309, https://doi.org/10.3390/vaccines8040654.

[13] Amanat, F. et al. A serological assay to detect SARS-CoV-2 seroconversion in humans. Nat. Med. 26, 1033-1036, doi:10.1038/s41591-020-0913-5.(2020).

[14] Perez-Riba, A. & Itzhaki, L. S. A method for rapid high-throughput biophysical analysis of proteins. Sci. Rep. 7, 9071, doi:10.1038/s41598-017-08664-w.(2017). PMC5567296.

[15] Zhao, J. C., Zhao, Z. D., Wang, W. & Gao, X. M. Prokaryotic expression, refolding, and purification of fragment 450-650 of the spike protein of SARS-coronavirus. Protein Expr. Purif. 39, 169-174, doi:10.1016/j.pep.2004.10.004.(2005). PMC7129930.

[16] A.R. Marquez-Ipina, et al., Serological test to determine exposure to SARS-CoV-2: ELISA based on the receptor-binding domain of the spike protein (S-RBDN318-V510) expressed in Escherichia coli, Diagnostics 11 (2021), https://doi.org/10.3390/diagnostics11020271.

[17] T. Burki, Understanding variants of SARS-CoV-2, Lancet 397 (2021) 462, https://doi.org/10.1016/S0140-6736(21)00298-1.

[18] Klumpp-Thomas, C. et al. D614G spike variant does not alter IgG, IgM, or IgA spike seroassay performance. J. Infect. Dis., doi:10.1093/infdis/jiaa743(2020). PMC7798976.

[19] E.B. Hodcroft, et al., Emergence in Late 2020 of Multiple Lineages of SARS-CoV-2 Spike Protein Variants Affecting Amino Acid Position 677, 2021, https://doi.org/10.1101/2021.02.12.21251658 medRxiv, 2021.2002.2012.21251658.