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.

Virology 557 (2021) 15–22

Contents lists available at ScienceDirect

Virology

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

Expression, purification and immunological characterization of recombinant nucleocapsid protein fragment from SARS-CoV-2

Teodora Djukic [a,#], Maja Mladenovic [a,#], Dragana Stanic-Vucinic [a], Jelena Radosavljevic [a], Katarina Smiljanic [a], Ljiljana Sabljic [b], Marija Devic [b], Danica Cujic [b], Tamara Vasovic [a], Ana Simovic [a], Mirjana Radomirovic [a], Tanja Cirkovic Velickovic [a,c,d,e,*]

[a] University of Belgrade–Faculty of Chemistry, Department of Biochemistry, Centre of Excellence for Molecular Food Sciences, Belgrade, Serbia
[b] University of Belgrade, Institute for the Application of Nuclear Energy – INEP, Belgrade, Serbia
[c] Ghent University Global Campus, Incheon, South Korea
[d] Ghent University, Faculty of Bioscience Engineering, Ghent, Belgium
[e] Serbian Academy of Sciences and Arts, Belgrade, Serbia

ARTICLE INFO

Keywords:
Recombinant nucleocapsid protein
COVID-19
SARS-CoV-2
Prokaryotic expression
serological assay

ABSTRACT

Serological testing is important method for diagnosis of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection. Nucleocapsid (N) protein is the most abundant virus derived protein and strong immunogen. We aimed to find its efficient, low-cost production.

SARS-CoV-2 recombinant fragment of nucleocapsid protein (rfNP; 58–419 aa) was expressed in *E. coli* in soluble form, purified and characterized biochemically and immunologically.

Purified rfNP has secondary structure of full-length recombinant N protein, with high percentage of disordered structure (34.2%) and of β-sheet (40.7%). rfNP was tested in immunoblot using sera of COVID-19 convalescent patients. ELISA was optimized with sera of RT-PCR confirmed positive symptomatic patients and healthy individuals. IgG detection sensitivity was 96% (47/50) and specificity 97% (67/68), while IgM detection was slightly lower (94% and 96.5%, respectively).

Cost-effective approach for soluble recombinant N protein fragment production was developed, with reliable IgG and IgM antibodies detection of SARS-CoV-2 infection.

1. Introduction

Coronavirus disease 2019 (COVID-19) was first reported from China in December 2019, and on March 12, 2020 outbreak has been classified as a global pandemic, still rapidly spreading and posing a great threat to global public health. Whole-genome sequencing results showed that the causative agent was a novel coronavirus, initially named 2019- nCoV by the World Health Organization (WHO) (Wu et al., 2020; Zhou et al., 2020; Zhu et al., 2020). Later, it is officially designated as SARS-CoV-2 by the International Committee on Taxonomy of Viruses (ICTV) and since recently, suggestion for a distinct name was proposed, human coronavirus 2019 (HCoV-19) (Gorbalenya et al., 2020; H. wei Jiang et al., 2020). Similar to SARS coronavirus (SARS-CoV-1), HCoV-19 can cause severe respiratory illness and significant mortality among those over 60 years old with chronic conditions.

In addition to worldwide used nucleic acid−based tests for detection of the virus during acute disease, for determination of the real infection rate and infection fatality rate in a population serosurveys are necessary. Serological assays are needed not only for these serosurveys, but also for identification of individuals who were infected (severe, mild, and asymptomatic) and who are potentially immune, as well as for identification of potential plasma donors. Beside, serological assays could be used for qualitative and quantitative characterization of the immune response to the virus (Stadlbauer et al., 2020). For development of a reliable serological assay of great importance is preparation of suitable SARS-CoV-2 antigens. The best candidates for antigens are structural SARS-CoV-2 proteins, spike protein (S protein), envelope protein (E), membrane protein (M), and nucleocapsid protein (N protein), and their fragments, obtained as recombinant proteins. The N protein is a 419-amino-acid alkaline protein with a short lysine-rich region,

* Corresponding author. University of Belgrade-Faculty of Chemistry, Belgrade, Serbia,
E-mail address: tcirkov@chem.bg.ac.rs (T. Cirkovic Velickovic).
# These authors equally contributed to the manuscript.

https://doi.org/10.1016/j.virol.2021.01.004
Received 28 September 2020; Accepted 11 January 2021
Available online 9 February 2021

suggested as the nuclear localization signal. It plays an important role in the process of virus particle assembly by enveloping the entire genomic RNA (Marra et al., 2003). It is the most abundant virus derived-protein, relatively conservative in coronaviruses, and it is strong immunogen in several coronaviruses (Timani et al., 2004). Hence it is often used as antigen for serological assays and for raising antibodies for diagnostic applications. Moreover, antibody to the nucleocapsid protein of SARS-CoV-2 is more sensitive than spike protein antibody for detecting early infection (Burbelo et al., 2020). The SARS-CoV-2 N protein can be divided into five regions; a predicted intrinsically disordered N-terminal arm (1–40 aa), N-terminal domain (NTD, e.g. an RNA binding domain, 41–186 aa), a predicted disordered central linker (187–257 aa), C-terminal domain (CTD, e.g. a dimerization domain, 258–361 aa), and a predicted disordered C-terminal domain (362–419 aa) (Cubuk et al., 2020). For sensitive and reliable serological assay it is necessary to produce SARS-CoV-2 N antigens. Protein expression in prokaryotic systems, such as *E. coli*, is cost effective way to rapidly provide high quantities of recombinant protein. In contrast to highly glycosylated S protein and its fragments, requiring eukaryotic expression, N protein of SARS-CoV-1 have shown to be successfully expressed in *E. coli* (Maache et al., 2006; Pei et al., 2005; Timani et al., 2004; Wu et al., 2004; Zuo et al., 2005). Although there are several studies using *E. coli* expressed HCoV-19 N protein there is no study presenting both structural and immunochemical characterization, using sera of COVID-19 patients, of recombinant SARS-CoV-2 N protein obtained in *E. coli* (Ye et al., 2020; Zeng et al., 2020; Zhang et al., 2020).

In this study, the recombinant SARS-CoV-2 N protein fragment (rfNP; residues from 58 to 419) was expressed in *E. coli* and purified to homogeneity. The purified rfNP was characterized by CD spectrometry and mass spectrometry, followed by its evaluation by immunoblot and ELISA using sera of SARS-CoV-2 patients.

2. Material and methods

2.1. Material

E. coli host strains BL21(DE3), were obtained from Novagen (Wisconsin, USA). Synthesis of construct (based on sequence from protein data base, UniProt ID ID P0DTC9) cloned in *Nde*I and *Xho*I sites of plasmid pET20b, was ordered from Genscript (Piscataway, New Jersey, United States). Agar, tryptone and yeast extract were purchased from Torlak (Institute for virology, vaccines and serums, Belgrade, Serbia). Imidazole, 5-Bromo-4-chloro-3-indolyl phosphate (BCIP) and nitro blue tetrazolium (NBT) were purchased from Carl Roth, (Germany) isopropyl-D-thiogalactopyranoside (IPTG), ampicillin-sodium salt and Tris – base were purchased from Fisher Scientific (Loughborough, UK). Immobilized metal ion affinity chromatography (IMAC) resin, Ni Sepharose 6 fast Flow, and SP Sepharose were purchased from GE Healthcare (Uppsala, Sweden). 6x-His Tag Mouse anti-Tag, alkaline phosphatase (AP) monoclonal antibodies were purchased from Invitrogen (Carlsbad, California, United States). Coomassie Brilliant Blue R-250 (CBB) from Sigma (St Louis, MO, USA).

2.2. Patient serum samples

All study participants and serum donors have given written informed consent for the use of their serum for ELISA SARS-CoV-2 antibodies assay development. For Western blot serum samples were collected from subjects recovered from COVID-19 admitted to INEP laboratory at the time of sampling nasopharyngeal swabs for the second negative result confirmation by use of real-time polymerase chain reaction (RT- PCR). As negative control, sera from healthy volunteers taken from INEP sera bank were used (collected during 2018–2019). For ELISA assay a total of 118 human serum samples were analyzed. Blood samples positive for COVID-19 (n = 50) were collected from symptomatic patients, considered positive according to the results of the quantitative RT-PCR. Presence of SARS-CoV-2 antibodies in these sera was confirmed by ELISA SARS-CoV-2 IgG and IgM (INEP, Belgrade, Serbia). Another 68 sera samples were taken from healthy persons from the INEP sera bank.

2.3. Expression of SARS -CoV-2 rfNP

The recombinant plasmid was transformed into *E. coli* strain BL21 (DE3), with 40 ng of plasmid being mixed with 50 μL of competent cells. Mixture was heat shocked for 30 s at 42 °C. After heat shock, super optimal medium was added and cells were left to recover for 1 h (37 °C). Cells were plated on a solid Lauria Bertani broth (LB) media containing ampicillin and left to grow overnight (16h). Positive bacterial colonies were selected and cultured in LB liquid medium, containing ampicillin, overnight at 37 °C and shaking (220 rpm). The overnight cultures were transferred in 250 mL of fresh liquid LB medium containing ampicillin to permit exponential growth. When the optical density (OD_{600}) reached 0.700, protein expression was induced by addition of 0.4 mM IPTG overnight at two different temperatures (25 °C and 37 °C). Culture was centrifuged (10 min at 3000 rpm). Cells were resuspended in the 15 mL of lysis buffer (20 mM Tris, 500 mM NaCl at pH 8) and lysed by sonication. Soluble fraction of lysate was separated by centrifugation (30 min at 3000 rpm) and purified by IMAC (Ni- Sepharose).

2.4. SDS-PAGE and western blot

Soluble fraction of lysate was separated by sodium dodecyl sulfate–polyacrylamide gel (SDS-PAGE) on 14% gel in reducing conditions according to Laemmli protocol ((Cambridge), 1970). Separated proteins were transferred onto nitrocellulose membranes at 56 mA for 1 h. Membranes were incubated with 10 mL of Tween 20 Tris buffer saline (TTBS), pH 8, containing 1% bovine serum albumin (BSA) overnight at 4 °C. The expressed proteins were incubated with monoclonal anti-His-tag antibodies diluted in TTBS containing 0.5% BSA (dilution 1:2000). Primary antibodies were labeled with AP. After the incubation membranes were washed 2 times with TTBS and then with Tris buffer saline (TBS), pH 8. The binding patterns were visualized with BCIP and NBT as substrates.

2.5. Purification of SARS-CoV-2 rfNP by IMAC

Soluble fraction of lysate (15 mL) was run through a 2 mL Ni-Sepharose column, equilibrated with lysis buffer, using step elution (Tris, 500 mM NaCl, pH 8 containing 5, 50, 100, 200, 250, 300 and 500 mM imidazole). Fractions were analyzed on SDS – PAGE and 200 mM and 250 mM fractions were pooled and subjected to dialysis against 20 mM acetate buffer pH 7.2. Fractions eluted with 200 mM imidazole were then loaded onto 0.5 mL SP-Sepharose column equilibrated with 20 mM acetate buffer pH 7.2. Elution was done by increasing ionic strength, in steps with 100 mM, 200 mM, 300 mM, 500 mM and 1 M NaCl in equilibration buffer. The fractions containing pure rfNP were eluted with 500 mM NaCl, collected and dialyzed against 20 mM phosphate buffer pH 7.2. Purified recombinant rfNP was stored, in 20 mM phosphate buffer, containing 320 mM NaCl and 5% glycerol, was stored at −20 °C.

2.6. Determination of rfNP concentration

The concentration of rfNp was determined spectrophotometrically at 280 nm. The extinction coefficient for proteins under native conditions (ε) was calculated from the equation proposed by Pace et al.: $\varepsilon(\text{mL mg}-1\ \text{cm}-1) = (5500\ \text{nW} + 1490\text{nY} + 125\ \text{nC})$, where nW, nY and nC are the numbers of Trp, Tyr and Cys respectively per polypeptide chain and M is the molecular mass (Da). For rfNP, M = 40100 Da, nW = 4, nY = 11 and nC = 0. From which $\varepsilon = 38390$ mL mg−1 cm−1 (Pace et al., 1995).

2.7. CD spectra measurements and CD spectra analysis

CD spectra were recorded on a JASCO J-815 spectropolarimeter (JASCO, Tokyo, Japan) under stable temperature of 25 °C, with N protein concentration of 0.31 mg/mL. Before CD measurements sample was extensively dialyzed against 20 mM potassium phosphate buffer (pH 7.2) during 24 h at 4 °C. The spectra were collected over the wavelength range 180–260 nm for far UV. Each spectrum was acquired three times, and the results were averaged. Results were expressed as residue average molar ellipticity as follows: [h] = h/(10 n C d), where h is measured ellipticity, n is number of N protein amino acid residues, C is molar concentration of N protein sample, d is path length of the cell. The secondary structure content was determined by the CONTIN program using the CDPro software package (http://lamar.colostate.edu/~sreeram/CDPro/main.html). For the calculations, reference protein set SP37 was selected.

2.8. Mass spectrometry (MS) analysis of rfNP

2.8.1. Preparative SDS – PAGE and in-gel digestion

Preparative SDS-PAGE was carried out on 14% polyacrylamide gel under reducing conditions. A 20 μg of purified protein/cm of well was loaded. After SDS-PAGE gel staining, protein bands were excised, washed, reduced with dithiothreitol, and alkylated with iodoacetamide, followed by in gel trypsin-digestion, trypsin/sample ratio of 1:30 (w/w), as already described (Shevchenko et al., 2007).

2.8.2. Identification of proteins by Ultra performance liquid chromatography-electron spray ionization coupled to tandem mass spectrometry

Tryptic peptides from in-gel digestions, were chromatographically separated using an liquid chromatographic system of a quaternary Accela 600 pump, Accela auto sampler (Thermo Fisher Scientific Inc., Bremen, Germany) and analytical Hypersil GOLD column (50 × 2.1 mm, 1.9 μm particle size). The mobile phases were (A) water (MS-grade) with 0.1% formic acid and (B) acetonitrile (MS-grade) with 0.1% formic acid. The gradient program was as follows: 0–2 min 95% A, 2–18 min 95-50% A, 18–23 min 50-30% A, 23–26 min 30-5% A, 26–28 min 5% A, 28–28.1 min 5–95% A, 28.1–30 min 95% A, with flow rate of 0.3 mL/min. Injection volumes were between 10 and 30 μL (0.7–3 μg of protein). This UHPLC system was coupled with LTQ Orbitrap hybrid mass spectrometer (Thermo Fisher Scientific Inc., Bremen, Germany) equipped with heated electrospray ionization source. The identification of proteins was performed by PEAKS Suite X (Bioinformatics Solutions Inc., Canada). Signature spectra of fragmented parent ions, were searched using the PEAKS-DB and PEAKS-PTM algorithms against a hybridized database consisting of SARS-CoV-2 taxon 2697049, and Escherichia coli taxon 562, UniProtKB sequences, downloaded on June 01, 2020 from http://www.uniprot.org/, and the Max Quant contaminant database downloaded on March 01, 2019 from http://www.coxdocs.org/doku.php?id). More details are provided in Supplementary Material. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD023341 and 10.6019/PXD023341.

2.9. Immunological characterization of rfNP

2.9.1. Immunoblot

Purified recombinant SARS-CoV-2 nucleocapsid protein was separated on 14% SDS-PAGE under reducing conditions. Proteins were transferred on nitrocellulose membrane (Thermo Scientific, MA, USA). Membrane was incubated in blocking buffer (1% BSA in10 mM PBS pH 7.2, containing 150 mM NaCl and 0,05% Tween-20) for 1 h at room temperature. Membrane was cut and individual strips were incubated overnight on 4 °C with sera samples (diluted 1:50 in blocking buffer) from COVID-19 patients and healthy controls. For testing of nonspecific binding, membranes were incubated only in blocking buffer. Following washing in Tween phosphate buffer saline pH 7.2 (TPBS), membranes were incubated for 1 h at room temperature with sheep anti-human IgG/HRP or anti-human IgM/HRP conjugate (INEP, Belgrade, Serbia), diluted 1:1000 in blocking buffer. After washing in PBST, signal was developed using EnVision Detection Systems Peroxidase/DAB (DAKO, CA, USA).

2.9.2. ELISA assay

Sera previously tested in the ELISA SARS-CoV-2 IgG and IgM, INEP, Serbia (10 positive and 10 negative), were used for optimization of the new ELISA based on rfNP. The plates were coated with rfNP in different concentration (0.5 μg/ml, 1 μg/ml, 2 μg/ml, 3 μg/ml) in 0.1 M carbonate buffer pH 9.5 (100 μl/well). The plates were incubated at +4 °C for 18 h and washed three times with phosphate-buffered saline pH 7.6 (PBS) containing 0.05% Tween 20 (TPBS). The antigen coated plates were incubated with blocking solution with BSA (0.1 mg BSA in PBST, 100 μl/well) for 1 h at 37 °C. After washing, human sera diluted in blocking solution were added (1:50, 1:100, 1:200, 1:300, 1:400, 1:500; 100 μl/well) and plates were incubated for 1 h at 37 °C. After washing, the HRP-conjugated sheep anti-human IgG and anti-human IgM (INEP, Belgrade, Serbia) were diluted in blocking solution and added to the wells (1:1000, 1:2000, 1:5000, 1:8000; 100 μl/well) and then incubated for 1 h at 37 °C. After washing 4 times with washing buffer (PBST), color was allowed to develop for 10 min with chromogenic solution (3,3′,5,5′-tetramethylbenzidine (TMB)/hydrogen peroxide) (INEP, Belrade, Serbia). After stopping the reaction with stop solution (1 M H_2SO_4, INEP, Belgrade, Serbia), optical densities were measured at 450 nm using an ELISA reader (Wallac Multilabel Counter 1420, PerkinElmer, Italy).

3. Statistical analysis

In total 118 sera samples (68 negative and 50 positive) were used for preliminary ELISA assay validation. Statistical analyses were carried out using MedCalc version 10.4.0.0 (MedCalc Software, Ostend, Belgium) (Zweig and Campbell, 1993). Sensitivity was defined as the proportion of correctly identified COVID-19 positive patients who were initially positive by RT-qPCR SARS-CoV-2 determination in respiratory samples. Specificity was defined as the proportion of naive participants who classified as positive as analyzed by ELISA SARS-CoV-2 nucleocapsid specific antibodies.

4. Results

4.1. Expression of rfNP

SARS-CoV-2 N protein fragment, containing 6 x His residues, was expressed in *E. coli* (BL21) cells under various temperature conditions. SDS-PAGE of the total cell lysates (Fig. 1. A) shows that, after induction of expression, the band at about 40 kDa, corresponding to molecular mass of rfNP, is more intensive. This band is more intensive after induction at 37 °C than at 25 °C, implying more efficient expression, and expression was further done under that conditions. Western blot of cell lysate soluble fraction using anti-His antibodies (Fig. 1. B) demonstrated that 40 kDa band was specifically recognized by the antibodies, thus confirming expression of rfNP in soluble form.

4.2. Purification of rfNP

In the first purification step by IMAC, rfNP was eluted with 200 mM and 250 mM imidazole (Fig. 1. C). In order to obtain more purified frNP, additional purification by cation exchange chromatography was used. rfNP is highly basic protein with pI 9.98 and net charge of about 20 at pH 7. Therefore, chromatography on strong cation exchanger at pH 7.2 was applied as the second purification step, allowing only binding of highly basic proteins. The IMAC fraction with higher purity (eluated with 200

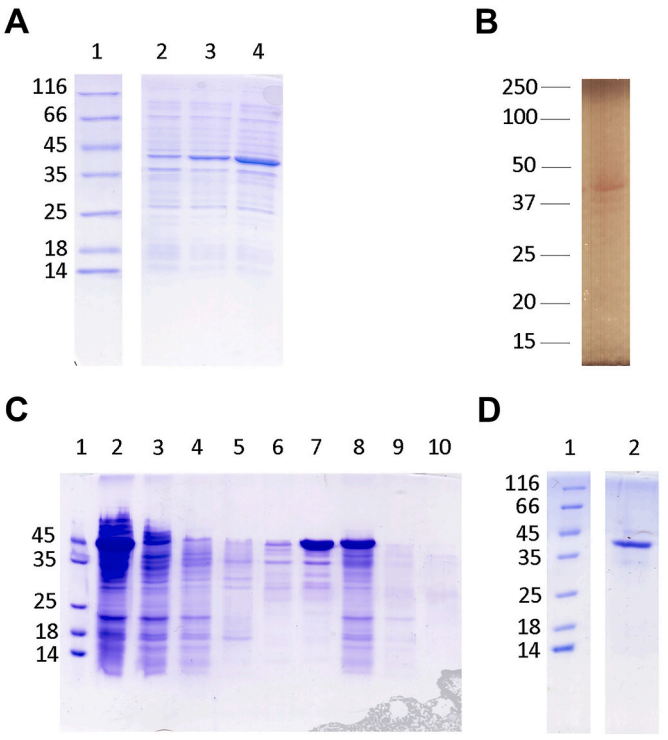


Fig. 1. Expression and purification of rfNP. (A) SDS-PAGE of total lysates of transfected *E. coli* BL21(DE3); lane 1 – MW markers, lane 2 – without induction, lane 3 after induction at 25 °C during 18 h, lane 4 - after induction at 37 °C during 18 h; (B) Western blot of *E. coli* expressed rfNP probed with anti-His antibodies; (C) frNP purification on Ni Sepharose 6 fast Flow; lane 1 – MW markers, lane 2 - soluble fraction of cell lysates; lane 3- flowthrough, lane 4 – unbound proteins; raction of unbound proteins; lanes 5–10 fractions eluated with imidayole: 50 mM (4),100 mM (5), 200 mM (6)), 250 mM (7), 300 mM (8) and 500 nm (9). (D) rfNP after second purification step by cation exchange chromatography on SP Sepharose; line 1 - MW markers, Line 2 - purified rfNP.

nM imidazole) was prepared for the second purification step, by dialysis against 20 mM acetate buffer pH 7.2. With the second chromatography step highly purified rfNP was obtained (Fig. 1. D).

4.3. Secondary structure determination

To determine secondary structure of recombinant N-protein CD spectral analysis was performed (Fig. 2. A) and the content of α-helix, β-sheet, β-turn and random coil were calculated (Fig. 2. B). CD spectral analysis showed high percentage of disordered structure (34.2%) and β-sheet (40.7%), whereas the least represented was α-helix (6.3%). CD spectra demonstrated that rfNP have completely preserved secondary structure of full-length N protein, with characteristic signals at 203 nm, 214 nm and 230 nm (Zeng et al., 2020). About 40% of N protein sequence have highly flexible almost disordered structure, reflecting in high percentage of random coil. Dominant β-sheet secondary structure originates from both NTD and CTD domains, while α-helix content originate from some short helices in CTD domain (Lo et al., 2013). Actually, all three intrinsically disordered regions (N-terminal arm, linker and C-terminal tail) are highly dynamic and contain regions of transient helicity that act as local binding interfaces for protein-protein or protein-RNA interactions (Cubuk et al., 2020).

4.4. Mass spectrometry analysis

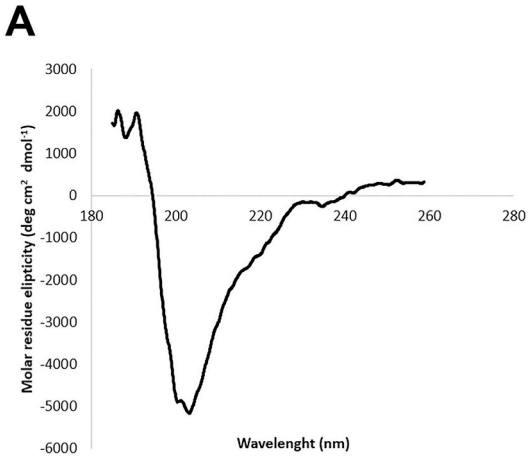
Purified rfNP was subjected to preparative SDS-PAGE and after in-gel digestion of excised band at 40 kDa, tandem mass spectrometry analyses were done. Tandem mass spectrometry identification of proteins in in-gel digested band, confirmed identity of rfNP with high scores and peptide coverage above 80% (Fig. 3, Table S1). Besides rfNP, the presence of 3 additional proteins from *E. coli*, was confidently detected within the same band (Table S1). From the value of areas under their ion extracted chromatographic curves (Table S1), it can be estimated than only up to 0,03% of the total band protein quantity belongs to host proteins, while rfNP share is well above 99,9%. The half of missing covered rfNP sequences were within the central linker domain (Fig. 3).

4.5. Immunoreactivity of purified rfNP tested by immunoblot

Interactions of purified rfNP with IgG and IgM antibodies from sera of COVID-19 convalescent patients and healthy individuals were tested (Fig. 4). For IgM antibodies, weak signal was obtained for majority tested samples, regardless of previous COVID-19 infection. On the other hand, samples from subjects with previous COVID-19 infection (samples 4,5,6 and 8) showed strong IgG reactivity with rfNP band. Also, faint band was observed in some COVID-19 negative samples for IgG antibodies. For both anti-human IgM and anti-human IgG antibodies nonspecific binding to rfNP was not observed.

4.6. Reactivity of rfNP in ELISA for SARS-CoV-2 antibodies detection

For the optimization of the new ELISA based of rfNP, 10 SARS-CoV-2 positive and 10 negative sera were used for the ELISA SARS-CoV-2 IgG and IgM testing. The biggest differences in OD values for negative and positive sera for both classes of antibodies were observed when using



Secondary structure	Percentage (%)
α-helix	6.3
β-sheet	40.7
β-turn	18.8
Random coil	34.2

Fig. 2. Structural characterization of rfNP. Far-UV CD spectra (A) and the secondary structure content (B) of purified rfNP.

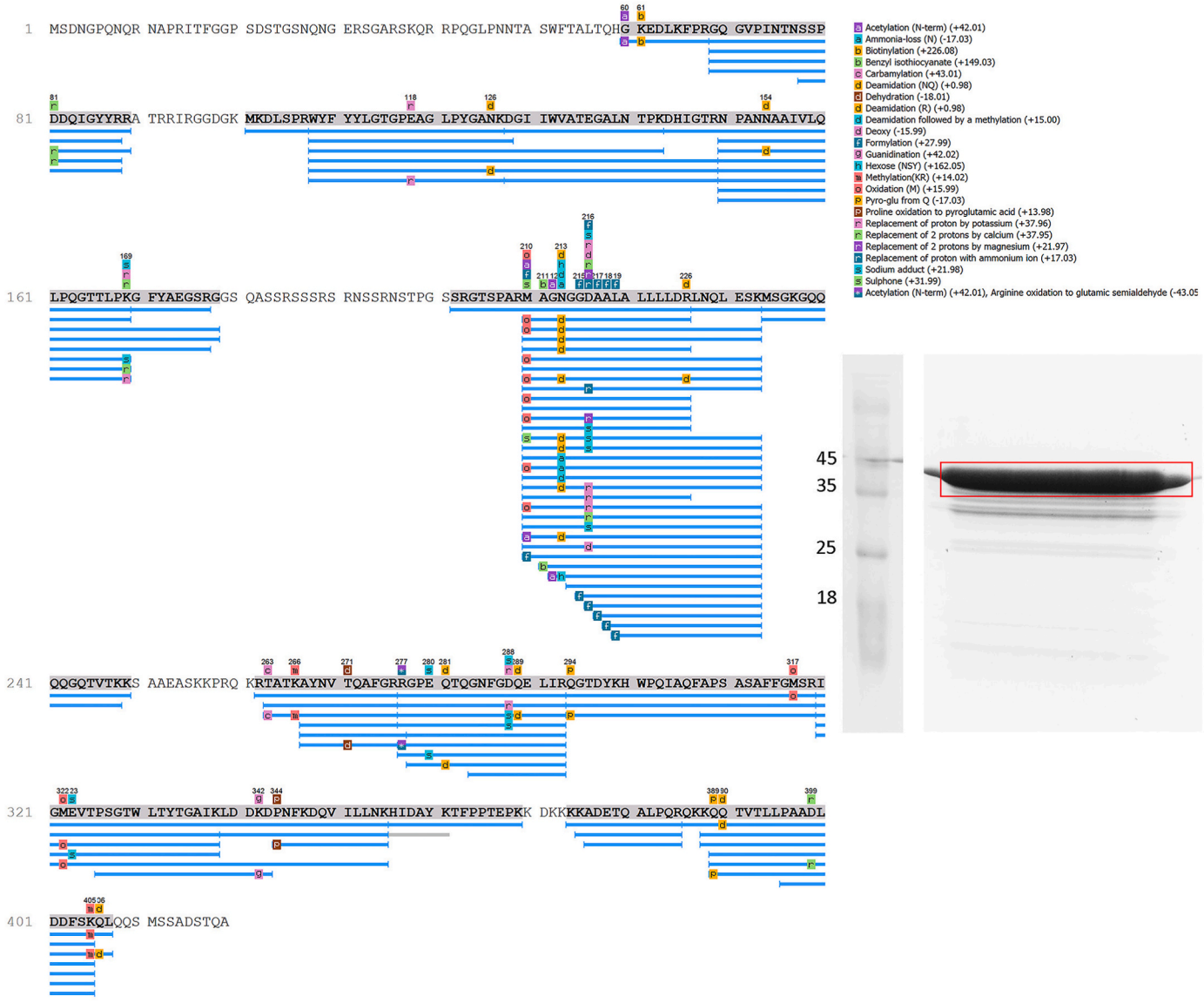


Fig. 3. Sequence coverage of N protein, obtained after mass spectrometry analysis of purified frNP; The inset presents preparative SDS PAGE of purified rfNP with marked 40 kDa band excised for in-gel digestion and MS analysis.

rfNP as the coating protein at a final concentration of 3 μg/ml (OD_{min} 0.08, OD_{max} 2.00). The optimal dilution of the test sera was 1:50. Using this optimal dilution of coating protein and sera, the optimal dilution of the HRP conjugated anti-human IgG was found to be 1:5000, and anti-human IgM was 1:8000. Under optimized conditions the average OD value for 10 negative sera samples ranged from 0.08 to 0.149 for IgG and 0.150–0.310 for IgM specific antibodies determination. The cut-off was determined to be 0.190 for IgG (Fig. 5 A) for IgM antibodies (Fig. 5. B). Based on the above-mentioned criteria, 67 samples from the 68 negative control sera were negative, and 47 samples from the 50 positive control sera were positive (OD 0.19–1.70) in the SARS-CoV-2 IgG ELISA. In the case of the detection of rfNP specific IgM antibodies, 64 sera from the group of negative control sera (n = 68) were negative in assay, and 44 sera from 50 SARS-CoV-2 positive samples were positive in SARS-CoV-2 IgM ELISA assay (OD 0.350–2.00). Distribution of the OD values obtained from 68 SARS-CoV-2 negative and 50 SARS-CoV-2 positive human sera using IgG and IgM ELISA are presented in Fig. 5. A and 5. B, respectively. According to the ROC (receiver operating characteristic) curve analysis, the best cut-off values were 0.190 OD for IgG detection (Fig. 5C) and using these values, sensitivity was 96% and specificity was 97%. In this test, the Area under the curve (AUC) was 0.983, which suggests that defined cut-offs are suitable for interpreting ELISA results and attainment of high precision in identifying the presence or absence of anti-SARS-CoV2 specific IgG antibodies in human sera (the maximum AUC is 1, which corresponds to a perfect classifier, whereas large AUC values indicate better classifier performance).

The best cut-off values for IgM detection (Fig. 5. D) was 0.350, with sensitivity of 94% and specificity of 95.6%. AUC was 0.977 indicating to acceptable degree of accuracy in detection of SARS-CoV-2 specific IgM antibodies in assay based on rfNP.

5. Discussion

In this study recombinant SARS-COV-2 N protein fragment was expressed in *E. coli*, structurally characterized and its potential for development of ELISA assay, for detection of SARS-CoV-2 specific IgM and IgG antibodies, was examined. Designed rfNP was successfully expressed in E. coli in soluble form. After purification by IMAC, rfNP highly positive charge at neutral pH was used for its further purification by cation exchange chromatography on strong cation exchanger at pH

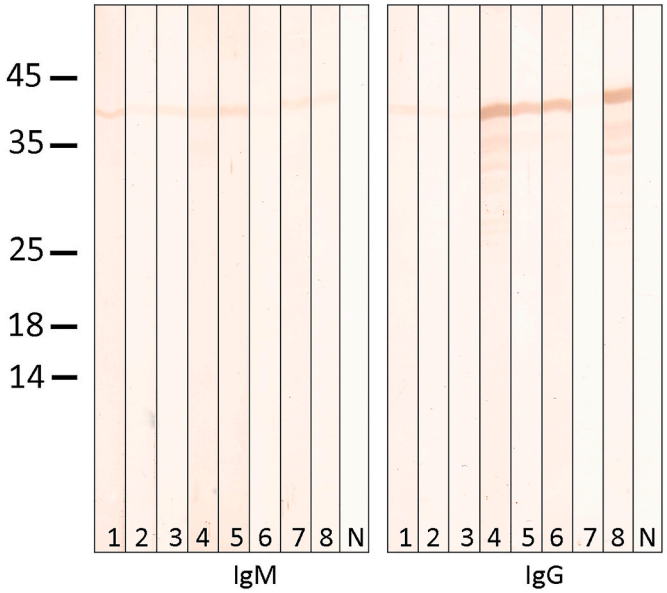


Fig. 4. Representative immunoblot of rfNP with sera from COVID-19 patients (samples 4,5,8) and healthy individuals (samples 1,2,3,7); N – nonspecific binding of anti-human IgM and IgG antibodies.

7.2. This is the first study exploiting N protein high pI for obtaining of high purity of recombinant SARS-COV-2 N protein or its fragments. Actually, almost all studies using E. coli expressed HCoV-19 N protein applied size exclusion chromatography as the second purification step (Kang et al., 2020; Ye et al., 2020; Zeng et al., 2020). CD spectrometry have shown that purified rfNP have secondary structure in accordance to full length recombinant N protein and its purification from soluble fraction of lysate contributed to maintaining its secondary structure (Zeng et al., 2020). In contrast, in the study of Zhang et al. (2020) recombinant N proteins (full-length, N-terminal and C-terminal fragments) expressed in *E. coli* were all insoluble precipitates, implying the loss of their native structure. Mass spectrometry analysis identified SARS-CoV-2 N protein in purified rfNP preparation with high sequence coverage (even 83%). In immunoblot purified rfNP have shown to be strong binder of IgG from sera COVID-19 convalescent patients, showing almost no binding to IgG from control healthy individuals. However, one control sera have shown low IgG binding due to the high conserved N protein sequences across the coronavirus species, which was also observed in the study of Jiang et al. (S. Jiang et al., 2020). It is well known that immunogenic proteins of closely related human coronaviruses can trigger cross-reactive antibodies in the host (Özçürümez et al., 2020). However, rfNP binding of IgM was low in non-COVID-19 and COVID-19 convalescent sera and at similar level. The low binding of IgM from non-COVID-19 sera was because of the cross-reactivity with other corona viruses, while low binding of COVID-19 convalescent sera is consequence of low IgM levels in sera of patients of at the convalescent stage due to IgM to IgG class-switch (Sun et al., 2020). It is to be noted that rfNP expressed in our study encompassed major IgG epitopes of N-protein, while it is lacking N-terminal domain that it is not immunologically relevant, thus preventing random non-specific reactions with circulating antibodies (Tilocca et al., 2020). ELISA test was optimized for SARS-CoV-2 IgG and IgM testing based on produced rfNP. In optimized assay, ROC curve analysis demonstrated sensitivity of 96% and specificity of 97%, indicating that, using rfNP, the likelihood of

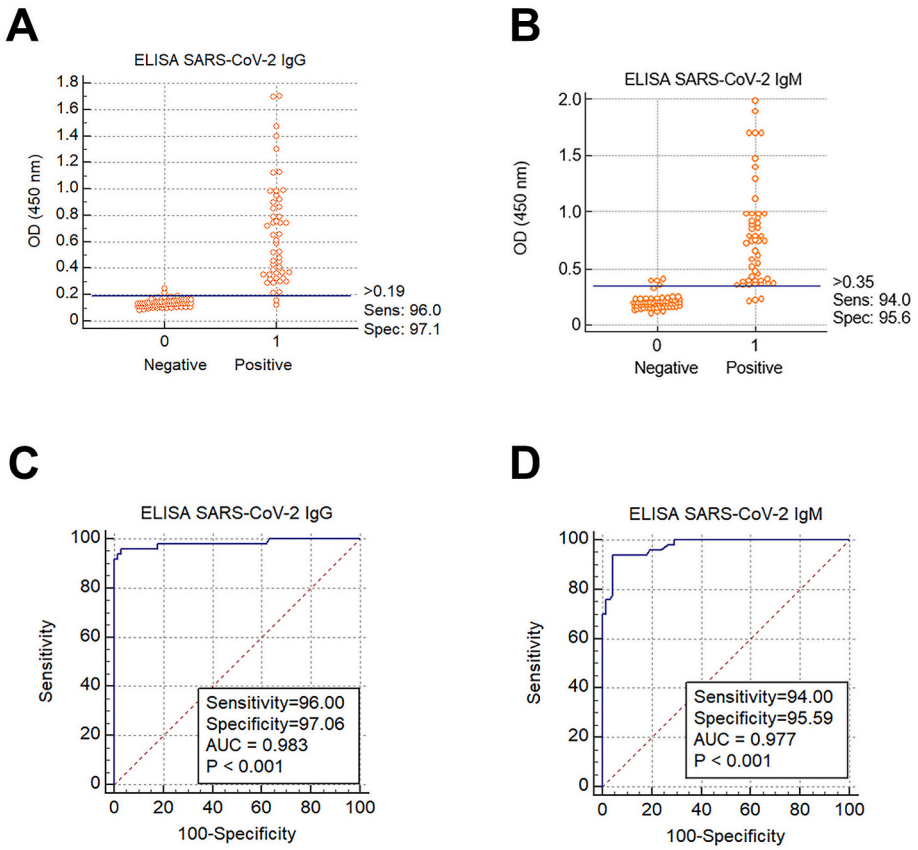


Fig. 5. Distribution of the OD values obtained from 68 SARS-CoV-2 negative and 50 SARS-CoV-2 positive human sera sample using rfNP based SARS-CoV-2 IgG (A) and IgM (B) indirect ELISA. The ROC curve built for 68 SARS-CoV-2 negative and 50 SARS-CoV-2 positive human sera analyzed by nucleocapsid based SARS-CoV-2 IgG (C) and IgM (D)ELISA.

false-positives is low, in spite of N protein cross-reactivity with other corona viruses. In ELISA optimized for rfNP-based IgM detection, acceptable, but lower sensitivity (94%) and specificity (95.6%) were obtained. The lower IgM sensitivity was due to its lower level or less frequent IgM response to SARS CoV-2, in comparison to IgG, also in active (symptomatic) patients (Long et al., 2020). The low level and infrequent IgM response was found to SARS CoV-2 as well as to SARS CoV-1 N protein (Chu et al., 2020; Leung et al., 2004). The lower specificity of IgM is consequence of higher cross-reactivity of corona-serogroups members for IgM. In comparison to IgG response, IgM response to N protein of SARS coronaviruses was reported to be less discriminatory between the SARS and non-SARS cohorts, for both SARS-CoV-1 and SARS-CoV-2 (Chu et al., 2020)(Leung et al., 2004).

6. Conclusion

In summary, we successfully achieved the expression and purification of the recombinant SARS-CoV-2 N protein fragment in *E. coli* and showed its potential as a useful option for development of the assays for serodiagnosis of SARS-CoV-2 infection. These results describe the production of N- protein fragment as an antigen for serological ELISA test, offering a safe cost-effective tool for detection of patients positive to SARS-CoV-2which could provide sensitive method for early detection of SARS- CoV-2 infection.

Ethics approval

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional guides on the care and use of laboratory animals. This study involves the use of existing human specimens and therefore is not considered research on human subjects for which informed consent is required, according to the guidelines in 'The rules of procedures of the Ethics Committee of INEP' (No. 02–832/1).

Consent for publication

Not applicable.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Funding

Ministry of Education, Science and Technological Development of Republic of Serbia, through Contract number: 451-03-68/2020-14/200168 and by the Science Fund of the Republic of Serbia GA #7542203, COVID-19-CAPSIDO. The authors thank the Government of the Republic of Serbian ad the Ministry of Education, Science and Technological Development of the Republic of Serbia for financial support.

Author contributions

TCV contributed to design the study, methodology, validation, resources, writing—review and editing, supervision, project administration and funding acquisition. DSV contributed to methodology (protein purification), validation, writing—review and editing. TD contributed to investigation (E. coli expression and purification), data curation, writing the original draft and figure preparation. MM contributed to investigation (*E. coli* protein expression and purification, formal CD spectral analysis), writing the original draft. KS contributed to methodology of MS analysis as well as software analysis, data curation, writing – review and editing. JR contributed to purification methodology and reviewing the original draft. LS contributed to immunological investigation. MD and DC contributed to methodology of immunological part of the study, data curation and writing the original draft. TV contributed to formal MS analysis and investigation (expression and purification). AS contributed to investigation (expression and purification). MR contributed to CD formal analysis.

Declaration of Competing interest

The authors declare no financial or commercial conflict of interest.

Acknowledgement

The authors acknowledge support for this research work that was funded by the Ministry of Education, Science and Technological Development of Republic of Serbia, through Contract number: 451-03-68/2020-14/200168; the Science Fund of Serbia, project Covid-19-CAPSIDO GA No.7542203, Serbian Academy of Sciences and Arts GA No. F-26. The authors thank the Government of the Republic of Serbian and the Ministry of Education, Science and Technological Development of the Republic of Serbia for financial support.

Abbreviations:

SARS-CoV-2	severe acute respiratory syndrome coronavirus 2
N	nucleocapsid
rfNP	recombinant fragment of nucleocapsid protein
aa	amino acid
COVID-19	coronavirus disease 2019
RT-PCR	real time-polymerase chain reaction
HCoV-19	human coronavirus
S	spike
E	envelope
M	membrane
ELISA	enzyme-linked immunosorbent assay
OD	optical density
CD	circular dichroism
BCIP	5-Bromo-4-chloro-3-indolyl phosphate
NBT	nitro blue tetrazolium
IPTG	isopropyl-D-thiogalactopyranoside
MAC	Immobilized Metal ion Affinity Chromatography
AP	alkaline phosphatase
CBB	Coomassie Brilliant Blue R-250
LB	Lauria Bertani
TBS	Tris buffer saline
TTBS	Tween 20 Tris buffered saline
TPBS	Tween phosphate buffer saline
HRP	horse radish peroxidase
SDS	**PAGE**, sodium dodecyl sulfate–polyacrylamide gel electrophoresis
BSA	bovine serum albumin
MS	mass spectrometry
ROC	receiver operating characteristic
AUC	area under the curve

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.virol.2021.01.004.

References

Burbelo, P.D., Riedo, F.X., Morishima, C., Rawlings, S., Smith, D., Das, S., Strich, J.R., Chertow, D.S., Davey, R.T., Cohen, J.I., 2020. Sensitivity in detection of antibodies to nucleocapsid and spike proteins of severe acute respiratory syndrome coronavirus 2 in patients with coronavirus disease 2019. J. Infect. Dis. 222, 206–213. https://doi.org/10.1093/infdis/jiaa273.

Chu, D.K.W., Pan, Y., Cheng, S.M.S., Hui, K.P.Y., Krishnan, P., Liu, Y., Ng, D.Y.M., Wan, C.K.C., Yang, P., Wang, Q., Peiris, M., Poon, L.L.M., 2020. Molecular diagnosis of a novel coronavirus (2019-nCoV) causing an outbreak of pneumonia. Clin. Chem. 7, 1–7. https://doi.org/10.1093/clinchem/hvaa029.

Cubuk, J., Alston, J.J., Incicco, J.J., Singh, S., Stuchell-Brereton, M.D., Ward, M.D., Zimmerman, M.I., Vithani, N., Griffith, D., Wagoner, J.A., Bowman, G.R., Hall, K.B., Soranno, A., Holehouse, A.S., 2020. The SARS-CoV-2 nucleocapsid protein is dynamic, disordered, and phase separates with RNA. bioRxiv 2, 2020. https://doi.org/10.1101/2020.06.17.158121.

Gorbalenya, A.E., Baker, S.C., Baric, R.S., de Groot, R.J., Drosten, C., Gulyaeva, A.A., Haagmans, B.L., Lauber, C., Leontovich, A.M., Neuman, B.W., Penzar, D., Perlman, S., Poon, L.L.M., Samborskiy, D.V., Sidorov, I.A., Sola, I., Ziebuhr, J., 2020. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. Nat. Microbiol. 5, 536–544. https://doi.org/10.1038/s41564-020-0695-z.

Jiang, H. wei, Li, Y., Zhang, H. nan, Wang, W., Yang, X., Qi, H., Li, H., Men, D., Zhou, J., Tao, S. ce, 2020. SARS-CoV-2 proteome microarray for global profiling of COVID-19 specific IgG and IgM responses. Nat. Commun. 11, 1–11. https://doi.org/10.1038/s41467-020-17488-8.

Jiang, S., Shi, Z., Shu, Y., Song, J., Gao, G.F., Tan, W., Guo, D., 2020. A distinct name is needed for the new coronavirus. Lancet 395, 949. https://doi.org/10.1016/S0140-6736(20)30419-0.

Kang, S., Yang, M., Hong, Z., Zhang, L., Huang, Z., Chen, X., He, S., Zhou, Ziliang, Zhou, Zhechong, Chen, Q., Yan, Y., Zhang, C., Shan, H., Chen, S., 2020. Crystal structure of SARS-CoV-2 nucleocapsid protein RNA binding domain reveals potential unique drug targeting sites. Acta Pharm. Sin. B 10 (7), 1228–3835. https://doi.org/10.1016/j.apsb.2020.04.009.

Leung, D.T.M., Tam, F.C.H., Chun, H.M., Chan, P.K.S., Cheung, J.L.K., Niu, H., Tam, J.S. L., Pak, L.L., 2004. Antibody response of patients with Severe Acute Respiratory Syndrome (SARS) targets the viral nucleocapsid. J. Infect. Dis. 190, 379–386. https://doi.org/10.1086/422040.

Lo, Y.S., Lin, S.Y., Wang, S.M., Wang, C.T., Chiu, Y.L., Huang, T.H., Hou, M.H., 2013. Oligomerization of the carboxyl terminal domain of the human coronavirus 229E nucleocapsid protein. FEBS Lett. 587, 120–127. https://doi.org/10.1016/j.febslet.2012.11.016.

Long, Q.X., Liu, B.Z., Deng, H.J., Wu, G.C., Deng, K., Chen, Y.K., Liao, P., Qiu, J.F., Lin, Y., Cai, X.F., Wang, D.Q., Hu, Y., Ren, J.H., Tang, N., Xu, Y.Y., Yu, L.H., Mo, Z., Gong, F., Zhang, X.L., Tian, W.G., Hu, L., Zhang, X.X., Xiang, J.L., Du, H.X., Liu, H. W., Lang, C.H., Luo, X.H., Wu, S.B., Cui, X.P., Zhou, Z., Zhu, M.M., Wang, J., Xue, C. J., Li, X.F., Wang, L., Li, Z.J., Wang, K., Niu, C.C., Yang, Q.J., Tang, X.J., Zhang, Y., Liu, X.M., Li, J.J., Zhang, D.C., Zhang, F., Liu, P., Yuan, J., Li, Q., Hu, J.L., Chen, J., Huang, A.L., 2020. Antibody responses to SARS-CoV-2 in patients with COVID-19. Nat. Med. 26, 845–848. https://doi.org/10.1038/s41591-020-0897-1.

Maache, M., Komurian-Pradel, F., Rajoharison, A., Perret, M., Berland, J.L., Pouzol, S., Bagnaud, A., Duverger, B., Xu, J., Osuna, A., Paranhos-Baccalà, G., 2006. False-positive results in a recombinant severe acute respiratory syndrome-associated coronavirus (SARS-CoV) nucleocapsid-based western blot assay were rectified by the use of two subunits (S1 and S2) of spike for detection of antibody to SARS-CoV. Clin. Vaccine Immunol. 13, 409–414. https://doi.org/10.1128/CVI.13.3.409-414.2006.

Marra, M.A., Jones, S.J.M., Astell, C.R., Holt, R.A., Brooks-wilson, A., Butterfield, Y.S.N., Khattra, J., Asano, J.K., Barber, S.A., Chan, S.Y., Cloutier, A., Coughlin, S.M., Freeman, D., Girn, N., Griffith, O.L., Leach, S.R., Mayo, M., Mcdonald, H., Montgomery, S.B., Pandoh, P.K., Petrescu, A.S., Robertson, A.G., Schein, J.E., Siddiqui, A., Smailus, D.E., Stott, J.M., Yang, G.S., Plummer, F., Andonov, A., Artsob, H., Bastien, N., Bernard, K., Booth, T.F., Bowness, D., Czub, M., Drebot, M., Fernando, L., Flick, R., Gray, M., Grolla, A., Jones, S., Feldmann, H., Meyers, A., Kabani, A., Li, Y., Normand, S., Stroher, U., Tipples, G.A., Tyler, S., Vogrig, R., Ward, D., Watson, B., Brunham, R.C., Krajden, M., Petric, M., Skowronski, D.M., Upton, C., Roper, R.L., 2003. The Genome Sequence of the SARS-Associated Coronavirus 300, 1399–1405.

Özçürümez, M.K., Ambrosch, A., Frey, O., Haselmann, V., Holdenrieder, S., Kiehntopf, M., Neumaier, M., Walter, M., Wenzel, F., Wölfel, R., Renz, H., 2020. SARS-CoV-2 antibody testing—questions to be asked. J. Allergy Clin. Immunol. 146, 35–43. https://doi.org/10.1016/j.jaci.2020.05.020.

Pace, C.N., Vajdos, F., Fee, L., Grimsley, G., Gray, T., 1995. How to measure and predict the molar absorption coefficient of a protein. Protein Sci. 4, 2411–2423. https://doi.org/10.1002/pro.5560041120.

Pei, H., Liu, J., Cheng, Y., Sun, C., Wang, C., Lu, Y., Ding, J., Zhou, J., Xiang, H., 2005. Expression of SARS-coronavirus nucleocapsid protein in Escherichia coli and Lactococcus lactis for serodiagnosis and mucosal vaccination. Appl. Microbiol. Biotechnol. 68, 220–227. https://doi.org/10.1007/s00253-004-1869-y.

Shevchenko, A., Tomas, H., Havliš, J., Olsen, J.V., Mann, M., 2007. In-gel digestion for mass spectrometric characterization of proteins and proteomes. Nat. Protoc. 1, 2856–2860. https://doi.org/10.1038/nprot.2006.468.

Stadlbauer, D., Amanat, F., Chromikova, V., Jiang, K., Strohmeier, S., Arunkumar, G.A., Tan, J., Bhavsar, D., Capuano, C., Kirkpatrick, E., Meade, P., Brito, R.N., Teo, C., McMahon, M., Simon, V., Krammer, F., 2020. SARS-CoV-2 Seroconversion in humans: a detailed protocol for a serological assay, antigen production, and test Setup. Curr. Protoc. Microbiol. 57, 1–15. https://doi.org/10.1002/cpmc.100.

Sun, B., Feng, Y., Mo, X., Zheng, P., Wang, Q., Li, P., Peng, P., Liu, X., Chen, Z., Huang, H., Zhang, Fan, Luo, W., Niu, X., Hu, P., Wang, L., Peng, H., Huang, Z., Feng, L., Li, Feng, Zhang, Fuchun, Li, Fang, Zhong, N., Chen, L., 2020. Kinetics of SARS-CoV-2 specific IgM and IgG responses in COVID-19 patients. Emerg. Microbes Infect. 9, 940–948. https://doi.org/10.1080/22221751.2020.1762515.

Tilocca, B., Soggiu, A., Sanguinetti, M., Musella, V., Britti, D., Bonizzi, L., Urbani, A., Roncada, P., 2020. Comparative computational analysis of SARS-CoV-2 nucleocapsid protein epitopes in taxonomically related coronaviruses. Microbes Infect 22, 188–194. https://doi.org/10.1016/j.micinf.2020.04.002.

Timani, K.A., Ye, Li, Ye, Linbai, Zhu, Y., Wu, Z., Gong, Z., 2004. Cloning, sequencing, expression, and purification of SARS-associated coronavirus nucleocapsid protein for serodiagnosis of SARS. J. Clin. Virol. 30, 309–312. https://doi.org/10.1016/j.jcv.2004.01.001.

Wu, X.D., Shang, B., Yang, R.F., Yu, H., Ma, Z.H., Shen, X., Ji, Y.Y., Lin, Y., Wu, Y., DiLin, G.M., Tian, L., Gan, X.Q., Yang, S., Jiang, W.H., Dai, E.H., Wang, X.Y., Jiang, H.L., Xie, Y.H., Zhu, X.L., Pei, G., Li, L., Wu, J.R., Sun, B., 2004. The spike protein of severe acute respiratory syndrome (SARS) is cleaved in virus infected Vero-E6 cells. Cell Res. 14, 400–406. https://doi.org/10.1038/sj.cr.7290240.

Wu, F., Zhao, S., Yu, B., Chen, Y.M., Wang, W., Song, Z.G., Hu, Y., Tao, Z.W., Tian, J.H., Pei, Y.Y., Yuan, M.L., Zhang, Y.L., Dai, F.H., Liu, Y., Wang, Q.M., Zheng, J.J., Xu, L., Holmes, E.C., Zhang, Y.Z., 2020. A new coronavirus associated with human respiratory disease in China. Nature 579, 265–269. https://doi.org/10.1038/s41586-020-2008-3.

Ye, Q., West, A.M.V., Silletti, S., Corbett, K.D., 2020. Architecture and self-assembly of the SARS-CoV -2 nucleocapsid protein. Protein Sci. 92093, 1–26. https://doi.org/10.1002/pro.3909.

Zeng, W., Liu, G., Ma, H., Zhao, D., Yang, Yunru, Liu, M., Mohammed, A., Zhao, C., Yang, Yun, Xie, J., Ding, C., Ma, X., Weng, J., Gao, Y., He, H., Jin, T., 2020. Biochemical characterization of SARS-CoV-2 nucleocapsid protein. Biochem. Biophys. Res. Commun. 527, 618–623. https://doi.org/10.1016/j.bbrc.2020.04.136.

Zhang, P., Gao, Q., Wang, T., Ke, Y., Mo, F., Jia, R., Liu, W., Liu, L., Zheng, S., Liu, Y., Li, L., Wang, Y., Xu, L., Hao, K., Yang, R., Li, S., Lin, C., Zhao, Y., 2020. Evaluation of Recombinant Nucleocapsid and Spike Proteins for Serological Diagnosis of Novel Coronavirus Disease 2019 (COVID-19). https://doi.org/10.1101/2020.03.17.20036954.

Zhou, P., Yang, X.Lou, Wang, X.G., Hu, B., Zhang, L., Zhang, W., Si, H.R., Zhu, Y., Li, B., Huang, C.L., Chen, H.D., Chen, J., Luo, Y., Guo, H., Jiang, R.Di, Liu, M.Q., Chen, Y., Shen, X.R., Wang, X., Zheng, X.S., Zhao, K., Chen, Q.J., Deng, F., Liu, L.L., Yan, B., Zhan, F.X., Wang, Y.Y., Xiao, G.F., Shi, Z.L., 2020. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature 579, 270–273. https://doi.org/10.1038/s41586-020-2012-7.

Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R., Niu, P., Zhan, F., Ma, X., Wang, D., Xu, W., Wu, G., Gao, G.F., Tan, W., 2020. A novel coronavirus from patients with pneumonia in China, 2019. N. Engl. J. Med. 382, 727–733. https://doi.org/10.1056/NEJMoa2001017.

Zuo, X., Mattern, M.R., Tan, R., Li, S., Hall, J., Sterner, D.E., Shoo, J., Tran, H., Lim, P., Sarafianos, S.G., Kazi, L., Navas-Martin, S., Weiss, S.R., Butt, T.R., 2005. Expression and purification of SARS coronavirus proteins using SUMO-fusions. Protein Expr. Purif. 42, 100–110. https://doi.org/10.1016/j.pep.2005.02.004.

Zweig, M.H., Campbell, G., 1993. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clin. Chem. 39, 561–577. https://doi.org/10.1093/clinchem/39.4.561.