Antigenic properties of the SARS-CoV-2 nucleoprotein are altered by the RNA admixture

Denis E. Kolesov[1], Maria V. Sinegubova[1], Irina V. Safenkova[2], Ivan I. Vorobiev[1], Nadezhda A. Orlova*[1]

[1] Laboratory of Mammalian Cell Bioengineering, Institute of Bioengineering, Research Center of Biotechnology of the Russian Academy of Sciences. Moscow, Russia.
[2] Laboratory of Immunobiochemistry, Institute of Biochemistry, Research Center of Biotechnology of the Russian Academy of Sciences. Moscow, Russia.

Corresponding Author:
Nadezhda Orlova[1]
Email address: nobiol@gmail.com

Abstract

Determining the presence of antibodies to the SARS-CoV-2 antigens is the best way to identify infected people, regardless of the development of symptoms of COVID-19. The nucleoprotein (NP) of the SARS-CoV-2 is an immunodominant antigen of the virus - anti-NP antibodies are detected in persons previously infected with the virus with the highest titers. Many test systems for detecting antibodies to SARS-CoV-2 contain NP or its fragments as antigen. The sensitivity and specificity of such test systems differ significantly, which can be explained by variations in the antigenic properties of NP caused by differences in the methods of its cultivation, isolation and purification. We investigated this effect for the E.coli-derived SARS-CoV-2 NP, obtained from the cytoplasm in the soluble form. We hypothesized that co-purified nucleic acids that form a strong complex with NP might negatively affect NP's antigenic properties. Therefore, we have established the NP purification method, which completely eliminates the RNA in the NP preparation. Two stages of RNA removal were used - treatment of the crude lysate of *E. coli* with RNase A and subsequent selective RNA elution with 2 M NaCl solution. The resulting NP without RNA has a significantly better signal-to-noise ratio when used as an ELISA antigen and tested with a control panel of serum samples with antibodies to SARS-CoV-2; therefore, it is preferable for *in vitro* diagnostic use. The same increase of the signal-to-noise ratio was detected for the free N-terminal domain of the NP. Complete removal of RNA complexed with NP during purification will significantly improve its antigenic properties, and the absence of RNA in NP preparations should be controlled during the production of this antigen.

Introduction

The coronavirus pandemic affects the life of the entire population of the globe, with 258 830 438 confirmed cases of COVID-19 for 26.11.2021. The emergence of new SARS-CoV-2 variants(Khan et al., 2021) escaping the immune defense and variants with increased contagiousness and the impossibility of global vaccination of the entire planet population suggest that the pandemic will continue for several more epidemic seasons.

41 There is still a need for mass and cost-effective production of tests. This problem is especially acute in
42 countries with underdeveloped economies. The widespread availability of serological tests will help
43 defeat the pandemic and return to normalcy.
44 Acute or past infection of the individual with the SARS-CoV-2 virus may be tested using various *in vitro*
45 diagnostic methods, mainly the RT-PCR or serological testing for viral antigens presence or anti-viral
46 antibodies (Kilic, Weissleder & Lee, 2020). Testing for anti-SARS-CoV-2 antibodies reliably indicates
47 the presence of anti-viral immunity, regardless of the presence or absence of symptoms of the disease.
48 The generalized data from the population-wide serological studies makes it possible to accurately assess
49 the transmission dynamics and the actual scale of the virus's spread (Dorigatti et al., 2021). During mass
50 vaccination and for quite a long time after it, serological tests will also play a critical role in assessing the
51 effectiveness and duration of the vaccine's action.
52 Potential antigens for SARS-CoV-2 serological tests were evaluated on their specificity against other
53 common human coronaviruses (HCoV-OC43, HCoV-HKU1, HCoV-229E, HcoV-NL63) (Tilocca et al.,
54 2020), (Dobaño et al., 2020) and it was found that among the four viral structural proteins (proteins S, E,
55 M, N), two most abundant proteins – S and N (NP) are sufficiently specific; they are most often used for
56 practical serological testing. The N-antigen is the most abundant proteinaceous component of the viral
57 particle; therefore, it can be assumed that N-antigen causes the appearance of the highest antibody titers
58 during the infection. Serological tests using the N antigen are expected to have the best sensitivity.
59 According to (Okba et al., 2020), ELISA with the S1 antigen was more specific in discriminating
60 COVID-19 patients and HCoV or MERS-CoV groups, and ELISA with the NP was more sensitive in
61 revealing the seropositivity of patients with mild COVID-19. In direct comparison, the S antigen gives
62 slightly higher ELISA values than the NP antigen for the COVID-19 patients(Chen et al., 2020).
63 Most of the approved and used to date SARS-CoV-2 vaccines do not elicit a humoral immune response to
64 NP (Martínez-Flores et al., 2021); therefore, serological tests based on both S and N antigens may help to
65 distinguish between vaccinated and/or people infected with SARS-CoV-2. The specificity and sensitivity
66 of NP-based serological tests should be as high as possible since S-antigen tests cannot be used as a
67 parallel or control method for detecting SARS-CoV-2 infection in a significant proportion of patients.
68 The nucleoprotein of beta coronaviruses (NP) is immunodominant and most abundantly expressed viral
69 protein in the cytoplasm of infected cells (Timani et al., 2004). It consists of two tightly folded domains,
70 N-terminal (NTD) and C-terminal (CTD or N2b), responsible for RNA binding and NP dimerization,
71 respectively, and the surrounding short unstructured regions N1a, N2a, and B / N3. Nucleoprotein is the
72 only cytoplasmic protein of coronaviruses that is purposefully captured by viral particles separating from
73 infected cells. The NP is synthesized in the cell cytoplasm, so it is not glycosylated. It contains a single
74 cluster of Ser/Thr phosphorylation sites in the N2a subdomain (Tugaeva et al., 2021), and, probably,
75 some NP molecules have acetylated Lys residues (Hatakeyama et al., 2021), so the recombinant NP from
76 *E. coli* will adequately represent the natural NP as the antigen. At the same time, the recombinant NP,
77 produced in mammalian or yeast cells as the secreted protein, contains multiple N- and O-glycan moieties
78 and, presumably, is partially shielded from binding antibodies (Supekar et al., 2020). *E. coli* derived NP,
79 its single domains, and deletion mutants are widely used in serological assays (Guo et al., 2020)(Yamaoka
80 et al., 2021).
81 NP is a highly basic protein prone to multimerization and non-specific binding of nucleic acids (Zeng et
82 al., 2020), and both the NTD and CTD contain highly positively charged regions (Chang et al., 2014).
83 Residual admixtures of host cell biopolymers and polymerization state of the NP could significantly
84 change depending on its purification from specific impurities; therefore, it is necessary to develop a

85   method for obtaining a SARS-CoV-2 NP preparation that provides the maximum efficiency of the
86   serological testing. Many publications (Nasrallah et al., 2021)(Tehrani et al., 2020) mention the highly
87   differing quality of commercially available NP-based tests. Still, they do not discuss specific impurities in
88   the antigen that can cause a decrease in the sensitivity of the tests or methods for specific purification of
89   these impurities (Liu et al., 2021).
90   We proposed that NP's antigenic properties may be significantly altered due to the presence of nucleic
91   acid impurities, which will force the NP to the multimeric state (Perdikari et al., 2020), partially block its
92   surface, and expel some of the NP-binding antibodies by the negatively charged patches. Methods of
93   nucleic acid removal from viral nucleoproteins are well known – treatment with nucleases and
94   dissociation of complexes with the high salt solutions (Damodaran & Kinsella, 1983). Here the present
95   the simple protocol of SARS-CoV-2 NP expression and purification and demonstrate the change in its
96   antigenic properties upon two-stage removal of nucleic acid contaminants.

Materials & Methods

Molecular cloning

99   The NP open reading frame flanked with the *Pci*I-*Hind*III sites was synthesized by Epoch Life Science,
100   Inc (Missouri City, TX, USA). The synthetic gene was cloned into the pHYP expression vector (Genbank
101   ID: MW187859), cut at the *Nco*I and *Hind*III sites. The codon usage of the ORF was optimized for *E. coli*
102   expression with the GENEius-Light software http://www.geneius.de/GENEius, and a C-terminal 10xHis
103   cluster was added to the ORF for further purification of the target protein by metal chelate
104   chromatography. The resulting expression plasmid, pHYP-NPC-10H (Genbank ID: MW187860,
105   Addgene #162789), was transformed into the *E. coli* BL21[DE3] cells.
106   The *Avr*II and *Nhe*I sites were introduced into the synthetic NP ORF using the Silent Mutator online tool
107   (https://www.molbiotools.com/silentmutator.html).
108   The N-terminal part of NP ORF was deleted using the *Avr*II and *Nhe*I restriction of the pHYP-NPC-10H
109   plasmid with subsequent ligation, resulting in the pHYP-NP-tail-C-10H plasmid, the protein coded
110   corresponds to residues 218-419 of the SARS-CoV-2 NP, protein code – CTD.
111   The C-terminal part was deleted using *Nhe*I-*Hind*III restriction of the pHyp-NP plasmid and exchanged to
112   the pair of annealed oligos AD-NPhead-NheF (5'-
113   CTAGCCCATCACCATCATCACCACCATCACCATCACTGATA-3'); AD-NPhead-HindR (5'-
114   AGCTTATCAGTGATGGTGATGGTGGTGATGATGGTGATGGG-3'), resulting in pHYP-NP-head-C-
115   10H plasmid, the protein coded corresponds to residues 1-220 of the SARS-CoV-2 NP, protein code –
116   NTD. Sequences of the expected proteins are shown in Fig. S1.

Bacterial culture

118   Expression plasmids were transformed to the BL21[DE3] cells. The resulting strains were grown in shake
119   flasks in the 2xYT medium with 30 mg/L kanamycin and 0.1% glucose at 37°C, induced with the 1 mM
120   isopropyl-β-D-thiogalactopyranoside at OD600 0.6-1 AU and cultivated at 30°C for 3 h.
121   Bacterial cells were collected by centrifugation, 1 g of wet cell paste obtained was resuspended in 10 ml
122   lysis buffer (20 mM sodium phosphate, pH 7.4; 10 mM imidazole, 10 µg/ml egg white lysozyme,
123   0.1% Triton X-100) and incubated 15 min on ice. Bacterial DNA was sheared with two sonication rounds
124   of 20 sec using a Bandelin Sonopuls ultrasonic homogenizer (BANDELIN electronic GmbH & Co. KG).
125   The bacterial lysates were cleared by centrifugation at $45,000 \times g$ for 10 min and then divided into two
126   parts, one of which was treated with 10 µg/ml RNAse A (Thermo Fischer Scientific) for 15 min on ice,
127   and the second one was left untreated (NP-RNA; NTD-RNA).

NP and NTD purification

129 Immobilized metal affinity chromatography was carried out at room temperature on the Akta Explorer
130 system and the Tricorn 5/5 column, packed with 1 ml of Ni-Sepharose resin (all Cytiva). The optical
131 density of the eluate was monitored and recorded using the instrument control software Unicorn (Cytiva).
132 The column was equilibrated with 20 mM sodium phosphate, pH 7.5; 500 mM NaCl, 10 mM imidazole
133 solution. Clarified bacterial lysates were applied to the column at 0.5 ml/min. The column was washed
134 with equilibration solution until a stable baseline at 1 ml/min flow (approximately 10 column volumes),
135 then with the equilibration solution with 50 mM imidazole at the same flow velocity for 10 column
136 volumes. The column was optionally washed with the equilibration solution with 2 M NaCl (NP, NP-
137 NaCl, NTD, NTD-NaCl samples) until the stable baseline. Target proteins were eluted with the
138 equilibration solution with 250 mM imidazole-HCl at 0.5 ml/min; main eluate peaks were collected as
139 single fractions. Ten minutes flow stop was performed during elution for the yield maximization. Purified
140 proteins were divided into aliquots, flash-frozen in liquid nitrogen, and stored frozen for further use.
141 Purified protein solutions were intentionally not desalted from NaCl and imidazole and used in further
142 experiments, due to rapid precipitation of the NP upon desalting to the PBS solution. Target protein
143 concentration was determined using the Bradford method with the Total protein kit (Sigma, USA).

144 **Nucleic acids concentration measurement**
145 The concentration of RNA and DNA in protein samples was measured using the Qubit fluorometer
146 (ThermoFischer Scientific) and Qubit RNA and Qubit DNA HS reagent kits.

147 **SDS-PAGE**
148 Proteins were analyzed by SDS–PAGE (12.5% acrylamide) in reducing conditions with the PageRuler
149 prestained marker (ThermoFisher Scientific), 5 μl/lane. The gels were stained with colloidal Coomassie
150 blue, de-stained with the deionized water until clear background, scanned using the flatbed scanner in the
151 transparent mode, saved as 16-bit grayscale images. The images were analyzed using the TotalLab TL120
152 software (Nonlinear Dynamics, UK).

153 **Analytical size-exclusion chromatography**
154 Size exclusion chromatography was performed according to (Sinegubova et al., 2021) with minor
155 changes. Mobile phase was 20 mM sodium phosphate pH 7.5, 100 mM imidazole-HCl, 150 mM or
156 300 mM or 2 M NaCl. Size exclusion chromatography with the multi-angle light scattering (MALS)
157 detection was performed as described in (Tugaeva et al., 2021) with minor changes – mobile phase was
158 20 mM sodium phosphate pH 7.5, 100 mM imidazole-HCl, 300 mM NaCl, flow 0.5 ml/min, column
159 Superdex 200 10/300 (Cytiva).

160 **Dynamic light scattering (DLS)**
161 The hydrodynamic diameter (Dh) of proteins was measured using DLS as described in (Safenkova et al.,
162 2016) with minor changes. Protein concentrations were in the 0.6-1 mg/ml range for all samples tested, 20
163 mM 20 mM sodium phosphate pH 7.5, 500 mM NaCl, 250 mM imidazole solution in all cases.

164 **Enzyme-linked immunosorbent assay (ELISA)**
165 Control pooled serum samples with established immunoreactivity in clinically approved tests were
166 obtained from Anti-SARS-CoV-2 Verification Panel for Serology Assays (NIBSC panel, National
167 Institute for Biological Standards and Control, Potters Bar Hertfordshire EN6 3QG, UK). The NIBSC
168 Panel is comprised of 37 samples, 23 positive samples, and 14 negative samples. We also used the
169 samples described in the article (Sinegubova et al., 2021), namely pre-COVID-19 normal human plasma
170 sample (Renam, Moscow, Russia) and control pooled serum samples obtained from patients with the
171 PCR-confirmed SARS-CoV-2 infection (Xema Co., Ltd, Moscow, Russia). All samples were tested in
172 triplicates if not stated otherwise.

173   Microtiter plate wells (Corning, USA.) were coated with 100 μl of 1μg/ml test antigen in PBS with 100
174   mM imidazole solution and incubated overnight at + 4ºC. The wells were washed three times with PBS–
175   0.02% Tween (PBST) and blocked for 1 h at 37ºC with 3% bovine serum albumin (BSA; Sigma) in PBS,
176   washed with PBST, and used immediately. Test sera were prediluted with 1% BSA-PBS, applied as serial
177   twofold dilutions in the 1:400 –1:25 600 range, and incubated for 1 h at 37°C. Wells were washed three
178   times with PBST, secondary anti-human IgG antibody-HRP conjugate (Xema Co., cat. T271X@1702)
179   was used at the 1:20 000 dilution. After 1 h at +37°C, wells were washed five times with the PBST and
180   100 μl of ready-to-use TMB solution (Xema Co.) was added to each well. The color was developed for
181   exactly 10 minutes at room temperature (+25±2°C). The reaction was stopped by addition of 100 μl of 5%
182   orthophosphoric acid per well. Absorbance at 450 nm was measured with a Multiskan EX plate reader
183   (Thermo Fischer Scientific).
184   **Mass spectrometry**
185   High-performance liquid chromatography-electrospray ionization-mass spectrometry (HPLC-ESI-MS)
186   analysis was performed with the Impact II QqTOF high-resolution mass-spectrometer (Bruker Daltonik,
187   Germany), the UHPLC (Bruker Daltonik), and the Nucleodur C4 300-5 ec 4.6*150 mm column
188   (Macherey-Nagel, Germany). Data acquisition conditions were as follows: flow 1 ml/min, split 1:10,
189   gradient elution from 2% to 98% B in 40 min (A: 0.1% formic acid in water, B: 0.1% formic acid in
190   acetonitrile), column temperature 40°C, 30-50 μg of sample per injection, ESI source in positive mode,
191   HV capillary at 4.5 kV, spray gas – nitrogen at 1.8 bar, dry gas – nitrogen at 8 L/min 220°C. Spectra were
192   processed with BioPharma Compass 3.1.1 (Bruker Daltonik, Germany).
193   **Statistical analysis**
194   ELISA data were analyzed by the Student's two-tailed unpaired t-test or one-way ANOVA with the post-
195   hoc Tukey-Kramer HSD test utilizing the GraphPad Prism (GraphPad Software, San Diego, CA, USA),
196   LibreOffice (https://www.libreoffice.org/) and the RStudio software (RStudio PBC, Boston, MA, USA).
197   All Student's and ANOVA p-values are presented with the raw OD readings data in the Supporting File.
198   Correlational analysis for ELISA data obtained for the NIBSC panel serum samples was performed for 5
199   semi-quantitative and quantitative assays (published results) and two experimentally performed assays
200   with full-length NP antigen variants. The analysis was performed for all 37 serum samples and the
201   COVID-19 positive sera subset. Pearson's r was calculated using the RStudio software. In the case of
202   Pearson's r 0–0.19 is regarded as very weak, 0.2–0.39 as weak, 0.40–0.59 as moderate, 0.6–0.79 as
203   strong, and 0.8–1 as very strong correlation. Scripts employed for the correlation analysis are deposited at
204   the GitHub website: https://github.com/d-kolesiko/ELISA.
205   # Results
206   **Expression and purification of the intact NP**
207   *E. coli* strain BL21 [DE3] was transformed with the plasmid pHYP-NPC-10H encoding the full-length
208   NP SARS-CoV-2 with a C-terminal 10xHis tag (Fig. 1 A). During induction at 30°C, abundant
209   expression of the target protein in a predominantly soluble form was observed (Fig. 1 B); therefore,
210   further optimization of the cultivation conditions was not carried out. The target protein was purified from
211   the supernatant of lyzed bacterial cells by the immobilized metal affinity chromatography (Fig. 1 C).
212   According to UV absorption data, the purified NP contained a significant amount of nucleic acid
213   admixture - 280:260 nm absorption ratio was 0.514, far below the expected 1.75 ratio of the pure protein.
214   According to the intercalating dye fluorescence intensity analysis performed with the Qubit fluorimeter
215   (Thermo Fischer Scientific), contaminating nucleic acids were found to be the RNA. RNA:protein mass
216   ratio was 25.0 % for the NP preparation, untreated with the RNAse A (NP-RNA protein sample). Lysis

217   and purification of the full-length NP were repeated with two additional RNA removal steps – treatment
218   of the crude cell lysate with the 10 mg/L RNAse A for 15 min on ice and column wash with the
219   equilibration solution, adjusted to the 2 M NaCl concentration. According to the chromatography traces
220   data (Fig. 1 D, E), 2 M NaCl solution was sufficient to dissociate the protein – RNA complexes and elute
221   the contaminating RNA. Subsequent elution of the target protein with the 250 mM imidazole solution
222   produced the elution peak with the 280/260 nm absorption ratio of 1.2, typical for the pure protein.
223   Imidazole elution of the target protein without the RNAse A treatment and 2 M NaCl wash produced a
224   peak with the 280/260 nm absorption ratio of 0.53, indicating the significant contamination with nucleic
225   acids. This observation is in line with the intercalating dye analysis of NP preparations. The stepwise
226   gradient of the NaCl concentration in the wash solution revealed that admixtures with the 280:260 nm
227   light adsorption ratio 1:2, typical for the nucleic acid, are eluted mostly with the 1 M and 1.5 M NaCl
228   solutions, making the 2 M NaCl solution sufficient for the complete removal of nucleic acids from the
229   RNase A – treated NP (Fig. S2).
230   It should be noted that the NP preparation, purified from RNA with both RNAse A treatment and high
231   salt column wash, was completely RNA-free; residual RNA level was below the detection limit of the
232   Quant-iT ™ RNA Assay Kit used, i.e., below the 0.05% (w/w). According to this measurement, the NP
233   was purified from the RNA at least 12-fold. Both samples of purified NP were tested for the DNA
234   admixture using the Quant-It DNA high sensitivity dye and were found to be DNA-free, residual DNA
235   level was below 12 ppm in both samples. We made two additional control NP preparations - one treated
236   with the RNAse A and not washed on-column with the 2 M NaCl, another – washed with 2 M NaCl and
237   untreated with the RNAse A, and used these preparations for ELISA. Nucleic acids in the NP preparations
238   visualized on the agarose gel (Fig. 1 F) turned out to be highly polymeric.
239   **Expression and purification of the NTD and CTD fragments**
240   NTD and CTD of the NP in the pHYP plasmid were transformed into the *E.coli* BL21[DE3] cells.
241   Producers were cultured in the same way as for the full-length NP. We did not detect the protein band
242   corresponding to the CTD upon induction even in the insoluble proteins fraction (Fig. S3). At the same
243   time, free NTD was overexpressed similarly to the full-length NP (Fig. 2 A) and was purified with or
244   without the RNA removal steps as was described above for the full-length NP (Fig. 2 B).
245   **NP oligomerization state analysis**
246   According to the ESI-MS data, the molecular mass of the purified full-length NP, treated with both RNase
247   A and 2 M NaCl wash, was 46866.94 Da, in good agreement with the expected molecular mass of the NP
248   without the N-terminal Met residue; expected average molecular mass 46865.95 Da (Fig. 2 C, Fig. S4).
249   Similar data were obtained for the NTD protein, treated with both RNase A and 2 M NaCl wash
250   (Fig. 2 D, Fig. S5), molecular mass obtained 24637.52 Da, expected molecular mass 24637.81 Da for the
251   NTD without N-terminal Met. The relative mobility of the NP on the SDS-PAGE corresponds to the
252   52.7 kDa, indicating the slight gel retardation typical for the His-tagged protein.
253   The oligomerization state of the full-length NP preparations was checked by the size-exclusion
254   chromatography in physiological (Fig. S6) and high-salt solutions (Fig. 2 E, F, Fig. S6). We observed the
255   anomalous retention time of the target protein on the Superdex-200 column in physiological conditions in
256   the presence and in the absence of contaminating RNA (Fig. S6). In the presence of 300 mM NaCl in the
257   mobile phase, the RNA-free NP was present mainly in the dimeric form according to the data of MALS
258   detector - 82.4 kDa (Fig. S7 A), but the retention time of the main peak corresponded to the 142 kDa with
259   the minor monomer peak (43.6 kDa) visible (Fig. 2 E). This anomalous behavior of the SARS-CoV-2 NP
260   was already described in other studies (Tugaeva et al., 2021) and may correspond to the complex shape of

261   the NP dimer. At the same time, the NP-RNA produced multiple minor peaks corresponding to various
262   oligomerization numbers, and the major peak was heavily retarded on the column and was not suitable for
263   molecular mass calculation using the direct MALS detection or retention time curve interpolation
264   (Fig. 2 F). Further increase in the ionic strength of the mobile phase to 2 M NaCl resulted in the
265   retardation of the major peaks for both NP samples due to the hydrophobic interaction of the NP and the
266   chromatography sorbent (Fig.S6 E, F). NTD protein, unlike the full-length NP, was present mostly in the
267   monomeric form, according to both MALS data and calibration curve interpolation (Fig S7 B, S10) in the
268   physiological conditions and in the presence of 300 mM NaCl (Fig. S10 A-D). Similar, but not the same,
269   resin retardation results were obtained for the NTD fragment – in the presence of 2 M NaCl the RNA-free
270   NTD was seen as multiple retarded peaks, including the normal monomeric form, and in the presence of
271   RNA only the retarded forms and small quantities of non-retarded highly polymeric form of the NTD
272   were seen; monomeric NTD is completely absent on the chromatography trace (Fig. S10 E, F).
273   The SEC analysis of NP and NTD variants shows that both proteins form various multimeric complexes,
274   apparently interacting with the chromatography resin at all NaCl concentrations used, so all protein
275   samples, including the partially purified NP and NTD variants, were subjected to the dynamic light
276   scattering analysis in solution.
277   Both completely purified NP and NTD proteins contained a mixture of small and large protein particles.
278   Nucleoprotein molecules were distributed nearly equally between 18 nm particles (radius corresponds to
279   the dimeric form) and large particles with D_h >1000 nm (Fig. S8 A, B), and most of the NTD molecules
280   are present in the presumably monomeric form, D_h 9.2 nm (Fig. S9 A, B). Ribonuclease A treatment
281   without the on-column 2 M NaCl wash was sufficient to remove the intermediate particles in the case of
282   NP protein; particle median sizes and mass distribution are very close for NP and NP RNase+, NaCl-
283   preparations (Fig S8 C, D). At the same time, the NTD monomer was not seen after RNase A treatment,
284   most NTD molecules remained in the highly aggregated form. Sodium chloride wash without the RNase
285   A treatment was also insufficient for generating small protein particles in the concentrated NP or NTD
286   solutions (Fig S8 E, F; S9 E, F).
287   For the NP-RNA preparation that had not undergone RNase or 2 M NaCl treatments, only large particles
288   were observed, forming the dominant population of particles with D_h >1000 nm (Fig. S8 G, H) and a high
289   % polydispersity (%Pd), equal to 70.4%. Oligomeric particles are not seen in this population at all. The
290   NTD-RNA protein sample also was found to contain mainly highly polymeric protein particles, with D_h
291   >1000 nm and %Pd 8.4, but there was a peak with an average particle size of 77 nm (%Pd 5.7 %) (Fig. S9
292   G, H). These results confirmed that only a combination of RNase A treatment and on-column 2M NaCl
293   wash is sufficient for depolymerization of the NP, which should maximize the surface accessibility for
294   binding with antibodies.
295   Oligomerization of the *E. coli*-derived NP from the closely related beta-coronaviruses SARS-CoV and
296   MERS-CoV has been previously investigated in several works: (He et al., 2004)(Cong et al., 2017), and it
297   was found that NP tends to change the oligomerization state in the presence of nucleic acids. Salt effects
298   on the SARS-CoV-2 NP polymerization were not studied before, but for the unrelated DNA-binding
299   protein of the herpesvirus VP19C, the formation of the protein-DNA complex and the subsequent
300   polymerization are inhibited by increase of the NaCl concentration to 1 M (Bera et al., 2014). Although a
301   complete understanding of the mechanism of interaction of nucleoprotein monomers with each other has
302   not been achieved, most of the data indicate that NP exists in the form of a multimer in the presence of
303   nucleic acid and in the dimeric form after extensive purification (Perdikari et al., 2020).
304   **ELISA**

305   Full-length NP preparations and NTD preparations were used as the immobilized antigen in the antibody
306   capture ELISA testing of the anti-SARS-CoV-2 antibody titers. Proteins were applied to the microplates
307   in exactly equal amounts; the protein concentration was determined in all cases using the Bradford
308   method, which is insensitive to the RNA admixture. Dose-response curves were obtained for positive and
309   negative pooled serum samples, and it can be clearly seen that RNA-free NP antigen gives a significantly
310   (Student's one-tailed paired t-test, all ELISA data analyzed using this test, if not stated otherwise) higher
311   signal at various dilutions of the COVID-19-positive pooled sera (Fig. 3 A).
312   A similar effect was retained for the NTD antigen (Fig. 3 B). At the same time, the pooled pre-COVID-19
313   sera, derived in-house from the lyophilized normal blood plasma, was significantly more reactive with the
314   RNA-complexed full-length NP (Fig. 3 C) than with the pure NP protein. The non-specific binding of the
315   antibodies in the pre-COVID pooled sera to the NTD antigen was very similar for pure NTD and NTD-
316   RNA. However, the NTD gave a much weaker ELISA signal than the full-length NP for the COVID-19-
317   positive pooled sera, resulting in a lower overall signal-to-noise ratio, determined as 10,9 for the NP and
318   8,9 for the NTD (Fig S11 F, S12 E). Further analysis was carried out for the full-length NP as the more
319   sensitive and specific antigen.
320   We also tested all partially purified NP and NTD preparations – treated with the RNase A alone or with
321   2 M NaCl column wash alone (Fig. S11, Fig. S12). Expectedly, partial antigen purification resulted in
322   diminished assay signal-to-noise ratios in both cases, mainly due to lower antibody binding to the RNA-
323   complexed antigens for the pooled PCR+ sera.
324   ELISA testing of the pure NP and NP-RNA antigen was performed for all serum samples from the
325   NIBSC control COVID-19 sera panel, tested at two dilutions – 1:400 and 1:1600 (Fig. 3 D, E) (p=0.008
326   and p=0.03 respectively, Welch Two Sample t-test, RStudio). Positivity indices were calculated as the
327   ratio of the optical density for the test sample and the mean optical density in negative samples wells plus
328   three times the standard deviation, according to the (Lardeux, Torrico & Aliaga, 2016) (OD/CO ratio).
329   For both dilutions tested, the statistically significant difference of the OD/CO ratio was detected for the
330   positive samples group – RNA-free NP antigen was significantly more reactive with antibodies from
331   positive samples. At the same time, the OD/CO ratio for negative samples was very similar for both sera
332   dilutions tested.
333   We tested the correlation of the OD/CO ratios obtained for NIBSC control panel, pure NP antigen, NP-
334   RNA antigen, and published signal strengths for several clinically approved serological tests (Supporting
335   Table 1). All datasets were strongly correlated with the Pearson's r> 0.8 (Fig. 4 B) (P-values are
336   approximated using the t or F distributions). Correlation analysis was performed again for the positive
337   subset of the NIBSC control panel, and it was found that both pure NP and NP-RNA antigens still
338   produce highly correlated results, Pearson's r> 0.8. At the same time, correlation of OD/CO for our test
339   results with the quantitative and semiquantitative results of few clinically approved test systems was
340   lower, 0.4 < r < 0.8, and correlation of experimentally obtained ELISA data and published data for S
341   antigen and its fragments was approximately the same to the correlation of experimentally obtained data
342   and NP-based approved tests data (Fig. 4 A) (P-values are approximated using the t or F distributions).
343   Generally, the correlation of quantitative results of various tests for positive samples from the NIBSC
344   control panel gives the Pearson's r in the 0.4 – 0.8 range. Similar Pearson's r values were found for three
345   various automated serological tests in the case of real-world datasets – serum samples from the Qatar
346   population (Nasrallah et al., 2021).
347   Thus, the RNA-free SARS-CoV-2 NP obtained as described in this article is expected to be more suitable
348   for serological assays than the conventionally purified NP containing the tightly bound RNA. Two

349 different techniques of RNA removal should be performed; neither alone is sufficient for the complete
350 purification of the NP antigen.
351

Discussion

353 It is generally believed that NP and S-protein of the SARS-CoV-2 are the most specific antigens for
354 serological testing. In some studies, the NP is claimed as the more specific antigen (Turbett et al., 2021),
355 other studies challenge the specificity of the NP-based automated assays by the cross-antigen testing of
356 the samples with borderline anti-NP signals (Rosadas et al., 2020) or by the direct comparative testing of
357 the NP-based and S-protein RBD-based assays (Zonneveld et al., 2021). Testing large amounts of serum
358 samples with both NP and S antigens allows decreasing the proportion of false determinations to less than
359 1% (Poljak et al., 2021).
360 A significant proportion of the approved anti-SARS-CoV-2 vaccines codes or contains only the S-protein,
361 so the vaccinated individuals are expected to be S-protein positive regardless of the COVID-19 disease. If
362 infected with the SARS-CoV-2, such individuals may still be reliably screened using the NP-based
363 serological tests. Although the general method of obtaining the recombinant beta-coronavirus NP is well
364 known (Timani et al., 2004), and the presence of the RNA admixture in the NP was also described
365 previously (Liu et al., 2021), alteration of its antigenic properties caused by the RNA impurity was not
366 reported yet. We have demonstrated that the NP as the ELISA antigen dramatically loses its
367 immunoreactivity upon binding the bacteria-derived RNA.
368 We have shown that the antigenic properties of the SARS-CoV-2 nucleoprotein change significantly upon
369 the complete removal of the associated RNA. In ELISA tests with the RNA-free NP antigen, we observed
370 a statistically significant increase in the observed positivity indices for control positive sera and virtually
371 no change in the nonspecific antigen binding for the negative control sera.
372 When testing a pooled serum sample obtained from PCR-confirmed patients and a pre-COVID-19 pooled
373 serum sample, it was also found that completely RNA-free NP is significantly less immunoreactive
374 against a known negative sample than NP with an admixture of RNA.
375 The change of the antigenic properties of NP also reproduced for its free N-terminal domain. Thus, the
376 partial loss of the antigenic properties of the RNA-bound NP cannot be explained by the rearrangement of
377 NP into multimeric complexes alone. Apparently, a significant part of the NP's surface is blocked by
378 nucleic acids from contact with antibodies. It can also be assumed that the actual level of extracellular
379 RNA in the studied serum samples might also influence the observed titer of antibodies to NP SARS-
380 CoV-2; a more detailed study of this issue may be the subject of further research.

Conclusions

382 We have demonstrated the importance of specific purification of NP preparations for serodiagnostic
383 testing from co-purified nucleic acids and identified one of the possible causes for the variation in the
384 results of different serological tests. We have developed a purification procedure that includes only the
385 linearly scalable operations and may be used without major changes to produce the desired quantities of
386 NP, suitable for serological testing. We believe that the imperfect specificity of NP-based serological tests
387 for SARS-CoV-2 could be associated with insufficient antigen purification from RNA. The specific
388 removal of nucleic acids during NP production would improve the selectivity of new and existing
389 serological tests.
390

Acknowledgements

392   We thank Alexander Ivanov (Institute of Molecular biology Russian Academy of Sciences,
393   Moscow, Russia) and Valentin Manuvera (Research Institute for Physico-Chemical Medicine,
394   Moscow, Russia) for valuable comments and discussions on NP properties; Dr. Yuri Lebedin,
395   Eugenia Kostrikina, and Xema Co., Ltd., Moscow, Russia for providing control sera samples.
396   We thank Nikolay N. Sluchanko (A.N. Bach Institute of Biochemistry, Federal Research Center
397   of Biotechnology of the Russian Academy of Sciences, Moscow, Russia) for the size-exclusion
398   chromatography with MALS detection analysis and valuable comments on the NP behavior in
399   solutions; Andrey M. Tsedilin (Institute of Biochemistry) for the ESI-MS analysis.
400   The measurements were carried out in the Shared-Access Equipment Centre "Industrial
401   Biotechnology" of the Research Center of Biotechnology of the Russian Academy of Sciences.
402   DNA sequencing was carried out in the inter-institutional Center for collective use "GENOME"
403   IMB RAS, organized with the support of the Russian Foundation of Basic Research.
404   The study was supported by the Russian Foundation for Basic Research (grant 20-58-55001) and
405   the Ministry of Science and Higher Education of the Russian Federation.
406

References

408

409   Rosadas C, Randell P, Khan M, McClure MO, Tedder RS. 2020. Testing for responses to the wrong
410 SARS-CoV-2 antigen? *The Lancet* 396:e23. DOI: 10.1016/S0140-6736(20)31830-4.
411   Chang CK, Hou MH, Chang CF, Hsiao CD, Huang TH. 2014. The SARS coronavirus nucleocapsid
412 protein – Forms and functions. *Antiviral Research* 103:39–50. DOI: 10.1016/J.ANTIVIRAL.2013.12.009.
413   Cong Y, Kriegenburg F, Haan CAM de, Reggiori F. 2017. Coronavirus nucleocapsid proteins
414 assemble constitutively in high molecular oligomers. *Scientific Reports 2017 7:1* 7:1–10. DOI:
415 10.1038/s41598-017-06062-w.
416   Hatakeyama D, Masuda T, Miki R, Ohtsuki S, Kuzuhara T. 2021. In-vitro acetylation of SARS-
417 CoV and SARS-CoV-2 nucleocapsid proteins by human PCAF and GCN5. *Biochemical and Biophysical*
418 *Research Communications* 557:273–279. DOI: 10.1016/j.bbrc.2021.03.173.
419   Dobaño C, Santano R, Jiménez A, Vidal M, Chi J, Melero NR, Popovic M, López-Aladid R,
420 Fernández-Barat L, Tortajada M, Carmona-Torre F, Reina G, Torres A, Mayor A, Carolis C, García-
421 Basteiro AL, Aguilar R, Moncunill G, Izquierdo L. 2020. Immunogenicity and crossreactivity of
422 antibodies to SARS-CoV-2 nucleocapsid protein. *medRxiv*:2020.12.19.20248551. DOI:
423 10.1101/2020.12.19.20248551.
424   Nasrallah GK, Dargham SR, Shurrab F, Al-Sadeq DW, Al-Jighefee H, Chemaitelly H, al Kanaani
425 Z, al Khal A, al Kuwari E, Coyle P, Jeremijenko A, Kaleeckal AH, Latif AN, Shaik RM, Rahim HFA,
426 Yassine HM, al Kuwari MG, Qotba H, al Romaihi HE, Tang P, Bertollini R, Al-Thani MH, Althani AA,
427 Abu-Raddad LJ. 2021. Analytic comparison between three high-throughput commercial SARS-CoV-2
428 antibody assays reveals minor discrepancies in a high-incidence population. *Scientific Reports* 11. DOI:
429 10.1038/s41598-021-91235-x.
430   Timani KA, Ye L, Ye L, Zhu Y, Wu Z, Gong Z. 2004. Cloning, sequencing, expression, and
431 purification of SARS-associated coronavirus nucleocapsid protein for serodiagnosis of SARS. *Journal of*
432 *Clinical Virology* 30:309–312. DOI: 10.1016/j.jcv.2004.01.001.

433   Tugaeva K v., Hawkins DEDP, Smith JLR, Bayfield OW, Ker DS, Sysoev AA, Klychnikov OI,
434   Antson AA, Sluchanko NN. 2021. The Mechanism of SARS-CoV-2 Nucleocapsid Protein Recognition
435   by the Human 14-3-3 Proteins :SARS-CoV-2 N association with host 14-3-3 proteins. *Journal of*
436   *Molecular Biology* 433. DOI: 10.1016/j.jmb.2021.166875.
437   Guo L, Ren L, Yang S, Xiao M, Chang D, Yang F, dela Cruz CS, Wang Y, Wu C, Xiao Y, Zhang
438   L, Han L, Dang S, Xu Y, Yang QW, Xu SY, Zhu HD, Xu YC, Jin Q, Sharma L, Wang L, Wang J. 2020.
439   Profiling early humoral response to diagnose novel coronavirus disease (COVID-19). *Clinical Infectious*
440   *Diseases* 71:778–785. DOI: 10.1093/cid/ciaa310.
441   Lardeux F, Torrico G, Aliaga C. 2016. Calculation of the ELISA's cut-off based on the change-
442   point analysis method for detection of Trypanosoma cruzi infection in Bolivian dogs in the absence of
443   controls. *Memórias do Instituto Oswaldo Cruz* 111:501. DOI: 10.1590/0074-02760160119.
444   Liu P, Zong Y, Jiang S, Jiao Y, Yu X. 2021. Development of a Nucleocapsid Protein-Based ELISA
445   for Detection of Human IgM and IgG Antibodies to SARS-CoV-2. *ACS Omega* 6:9667–9671. DOI:
446   10.1021/ACSOMEGA.1C00253.
447   Poljak M, Oštrbenk Valenčak A, Štamol T, Seme K. 2021. Head-to-head comparison of two rapid
448   high-throughput automated electrochemiluminescence immunoassays targeting total antibodies to the
449   SARS-CoV-2 nucleoprotein and spike protein receptor binding domain. *Journal of Clinical Virology* 137.
450   DOI: 10.1016/j.jcv.2021.104784.
451   Okba NMA, Müller MA, Li W, Wang C, GeurtsvanKessel CH, Corman VM, Lamers mM,
452   Sikkema RS, Bruin E de, Chandler FD, Yazdanpanah Y, Hingrat Q le, Descamps D, Houhou-Fidouh N,
453   Reusken CBEM, Bosch B-J, Drosten C, Koopmans MPG, Haagmans BL. 2020. Severe Acute Respiratory
454   Syndrome Coronavirus 2−Specific Antibody Responses in Coronavirus Disease Patients - Volume 26,
455   Number 7—July 2020 - Emerging Infectious Diseases journal - CDC. *Emerging Infectious Diseases*
456   26:1478–1488. DOI: 10.3201/EID2607.200841.
457   Perdikari TM, Murthy AC, Ryan VH, Watters S, Naik MT, Fawzi NL. 2020. SARS-CoV-2
458   nucleocapsid protein phase-separates with RNA and with human hnRNPs. *The EMBO Journal*
459   39:e106478. DOI: 10.15252/EMBJ.2020106478.
460   He R, Dobie F, Ballantine M, Leeson A, Li Y, Bastien N, Cutts T, Andonov A, Cao J, Booth TF,
461   Plummer FA, Tyler S, Baker L, Li X. 2004. Analysis of multimerization of the SARS coronavirus
462   nucleocapsid protein. *Biochemical and Biophysical Research Communications* 316:476–483. DOI:
463   10.1016/j.bbrc.2004.02.074.
464   Zonneveld R, Jurriaans S, van Gool T, Hofstra JJ, Hekker TAM, Defoer P, Broekhuizen-van
465   Haaften PE, Wentink-Bonnema EM, Boonkamp L, Teunissen CE, Heijboer AC, Martens F, de Bree G,
466   van Vugt M, van Houdt R, van Agtmael M, Algera AG, van Baarle F, Bax D, Beudel M, Bogaard HJ,
467   Bomers M, Bos L, Botta M, de Brabander J, Brouwer MC, de Bruin S, Bugiani M, Bulle E, Chouchane
468   O, Cloherty A, Elbers P, Fleuren L, Geerlings S, Geerts B, Geijtenbeek T, Girbes A, Goorhuis A,
469   Grobusch MP, Hafkamp F, Hagens L, Hamann J, Harris V, Hemke R, Hermans SM, Heunks L, Hollmann
470   MW, Horn J, Hovius JW, Koning R, van Mourik N, Nellen J, Paulus F, Peters E, van der Poll T, Preckel
471   B, Prins JM, Raasveld J, Reijnders T, Schinkel M, Schultz MJ, Schuurman A, Sigaloff K, Smit M, Stijnis
472   CS, Stilma W, Thoral P, Tsonas A, van der Valk M, Veelo D, Vlaar APJ, de Vries H, Wiersinga WJ,
473   Wouters D, Zwinderman AH, van de Beek D. 2021. Head-to-head validation of six immunoassays for
474   SARS-CoV-2 in hospitalized patients. *Journal of Clinical Virology* 139. DOI: 10.1016/j.jcv.2021.104821.
475   Turbett SE, Anahtar M, Dighe AS, Beltran WG, Miller T, Scott H, Durbin SM, Bharadwaj M,
476   Thomas J, Gogakos TS, Astudillo M, Lennerz J, Rosenberg ES, Branda JA. 2021. Evaluation of three

commercial SARS-CoV-2 serologic assays and their performance in two-test algorithms. *Journal of Clinical Microbiology* 59. DOI: 10.1128/JCM.01892-20.

Sinegubova MV., Orlova NA, Kovnir SV., Dayanova LK, Vorobiev II. 2021. High-level expression of the monomeric SARS-CoV-2 S protein RBD 320-537 in stably transfected CHO cells by the EEF1A1-based plasmid vector. *PLOS ONE* 16:e0242890. DOI: 10.1371/JOURNAL.PONE.0242890.

Supekar NT, Shajahan A, Gleinich AS, Rouhani D, Heiss C, Azadi P. 2020. SARS-CoV-2 Nucleocapsid protein is decorated with multiple N- and O-glycans. *bioRxiv*:2020.08.26.269043. DOI: 10.1101/2020.08.26.269043.

Kilic T, Weissleder R, Lee H. 2020. Molecular and Immunological Diagnostic Tests of COVID-19: Current Status and Challenges. *iScience* 23. DOI: 10.1016/j.isci.2020.101406.

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 and Infection* 22:188–194. DOI: 10.1016/J.MICINF.2020.04.002.

Perdikari TM, Murthy AC, Ryan VH, Watters S, Naik MT, Fawzi NL. 2020. SARS-CoV-2 nucleocapsid protein phase-separates with RNA and with human hnRNPs. *The EMBO Journal* 39. DOI: 10.15252/embj.2020106478.

Yamaoka Y, Jeremiah SS, Miyakawa K, Saji R, Nishii M, Takeuchi I, Ryo A. 2021. Whole Nucleocapsid Protein of Severe Acute Respiratory Syndrome Coronavirus 2 May Cause False-Positive Results in Serological Assays. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 72:1291–1292. DOI: 10.1093/cid/ciaa637.

Zeng W, Liu G, Ma H, Zhao D, Yang Y, Liu M, Mohammed A, Zhao C, Yang Y, Xie J, Ding C, Ma X, Weng J, Gao Y, He H, Jin T. 2020. Biochemical characterization of SARS-CoV-2 nucleocapsid protein. *Biochemical and Biophysical Research Communications* 527:618–623. DOI: 10.1016/J.BBRC.2020.04.136.

Bera A, Perkins EM, Zhu J, Zhu H, Desai P. 2014. DNA Binding and Condensation Properties of the Herpes Simplex Virus Type 1 Triplex Protein VP19C. *PLOS ONE* 9:e104640. DOI: 10.1371/JOURNAL.PONE.0104640.

Chen Y, Tong X, Li Y, Gu B, Yan J, Liu Y, Shen H, Huang R, Wu C. 2020. A comprehensive, longitudinal analysis of humoral responses specific to four recombinant antigens of SARS-CoV-2 in severe and non-severe COVID-19 patients. *PLOS Pathogens* 16:e1008796. DOI: 10.1371/JOURNAL.PPAT.1008796.

Damodaran S, Kinsella JE. 1983. Dissociation of nucleoprotein complexes by chaotropic salts. *FEBS Letters* 158:53–57. DOI: 10.1016/0014-5793(83)80675-9.

Dorigatti I, Lavezzo E, Manuto L, Ciavarella C, Pacenti M, Boldrin C, Cattai M, Saluzzo F, Franchin E, Del Vecchio C, Caldart F, Castelli G, Nicoletti M, Nieddu E, Salvadoretti E, Labella B, Fava L, Guglielmo S, Fascina M, Grazioli M, Alvisi G, Vanuzzo MC, Zupo T, Calandrin R, Lisi V, Rossi L, Castagliuolo I, Merigliano S, Unwin HJT, Plebani M, Padoan A, Brazzale AR, Toppo S, Ferguson NM, Donnelly CA, Crisanti A. 2021. SARS-CoV-2 antibody dynamics and transmission from community-wide serological testing in the Italian municipality of Vo'. *Nature Communications 2021 12:1* 12:1–11. DOI: 10.1038/s41467-021-24622-7.

Khan A, Khan T, Ali S, Aftab S, Wang Y, Qiankun W, Khan M, Suleman M, Ali S, Heng W, Ali SS, Wei DQ, Mohammad A. 2021. SARS-CoV-2 new variants: Characteristic features and impact on the efficacy of different vaccines. *Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie* 143. DOI: 10.1016/J.BIOPHA.2021.112176.

Liu PP, Zong Y, Jiang SP, Jiao YJ, Yu XJ. 2021. Development of a Nucleocapsid Protein-Based ELISA for Detection of Human IgM and IgG Antibodies to SARS-CoV-2. *ACS omega* 6:9667–9671. DOI: 10.1021/ACSOMEGA.1C00253.

524   Martínez-Flores D, Zepeda-Cervantes J, Cruz-Reséndiz A, Aguirre-Sampieri S, Sampieri A, Vaca L.
525   2021. SARS-CoV-2 Vaccines Based on the Spike Glycoprotein and Implications of New Viral
526   Variants. *Frontiers in Immunology* 12:2774. DOI: 10.3389/FIMMU.2021.701501/BIBTEX.
527   Nasrallah GK, Dargham SR, Shurrab F, Al-Sadeq DW, Al-Jighefee H, Chemaitelly H, Al Kanaani Z, Al
528   Khal A, Al Kuwari E, Coyle P, Jeremijenko A, Kaleeckal AH, Latif AN, Shaik RM, Rahim HFA,
529   Yassine HM, Al Kuwari MG, Qotba H, Al Romaihi HE, Tang P, Bertollini R, Al-Thani MH,
530   Althani AA, Abu-Raddad LJ. 2021. Analytic comparison between three high-throughput
531   commercial SARS-CoV-2 antibody assays reveals minor discrepancies in a high-incidence
532   population. *Scientific Reports* 11. DOI: 10.1038/s41598-021-91235-x.
533   Safenkova I V., Slutskaya ES, Panferov VG, Zherdev A V., Dzantiev BB. 2016. Complex analysis of
534   concentrated antibody-gold nanoparticle conjugates' mixtures using asymmetric flow field-flow
535   fractionation. *Journal of Chromatography A* 1477:56–63. DOI: 10.1016/J.CHROMA.2016.11.040.
536   Tehrani ZR, Saadat S, Saleh E, Ouyang X, Constantine N, DeVico AL, Harris AD, Lewis GK, Kottilil S,
537   Sajadi MM. 2020. Performance of nucleocapsid and spike-based SARS-CoV-2 serologic assays.
538   *PLOS ONE* 15:e0237828. DOI: 10.1371/JOURNAL.PONE.0237828.
539   Timani KA, Ye L, Ye L, Zhu Y, Wu Z, Gong Z. 2004. Cloning, sequencing, expression, and purification
540   of SARS-associated coronavirus nucleocapsid protein for serodiagnosis of SARS. *Journal of
541   Clinical Virology* 30:309–312. DOI: 10.1016/J.JCV.2004.01.001.
542   Tugaeva K V., Hawkins DEDP, Smith JLR, Bayfield OW, Ker DS, Sysoev AA, Klychnikov OI, Antson
543   AA, Sluchanko NN. 2021. The Mechanism of SARS-CoV-2 Nucleocapsid Protein Recognition by
544   the Human 14-3-3 Proteins :SARS-CoV-2 N association with host 14-3-3 proteins. *Journal of
545   Molecular Biology* 433. DOI: 10.1016/j.jmb.2021.166875.
546

Figure 1

Expression and purification of the full-length SARS-CoV-2 nucleoprotein.

(A) - Map of the pHYP-NPC-10H plasmid. 10xHIS – C-terminal decahistidine tag; KanR2 – kanamycin A resistance gene; HS - hok/sok post-segregation killing locus; lac I – gene of the LacI repressor protein; T7 prom - promoter for bacteriophage T7 RNA polymerase. (B) – SDS-PAGE analysis of the full-length NP expression. Induction with 1 mM IPTG. (C) – SDS-PAGE analysis of the NP purification in the presence or absence of the RNase A. Target protein position is marked with the red arrow. "Non-binding" – column flowthrough fraction. (D), (E) - IMAC chromatography traces for the NP purification in the presence or absence of RNase A treatment and 2 M NaCl solution wash. Eluate absorbance at 280 nm is in blue, absorbance at 260 nm – in red. (F) – agarose gel analysis of purified proteins, ethidium bromide staining, double-strand DNA marker.
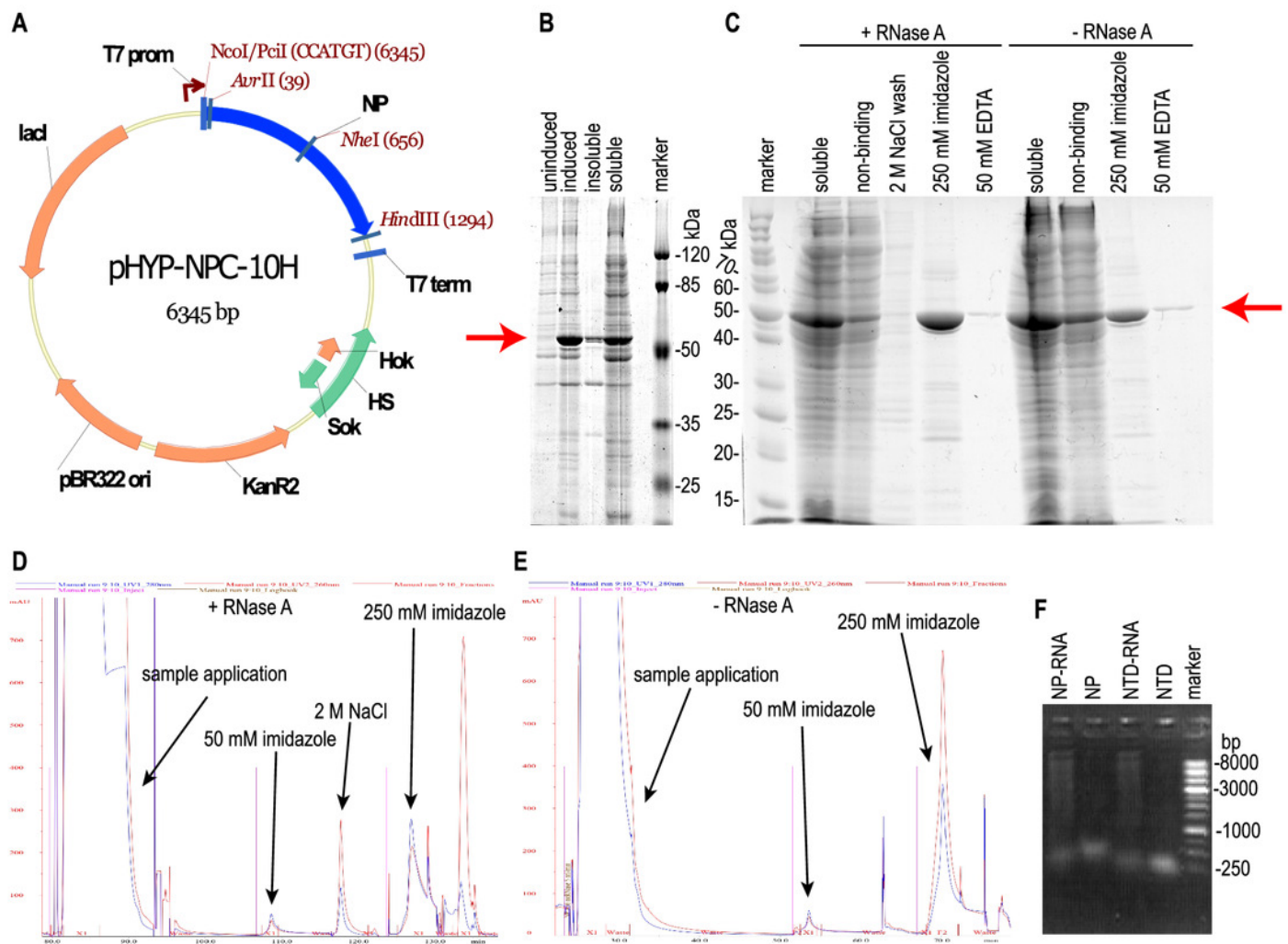
Figure 2

Expression and purification of the N-terminal domain of the nucleoprotein, analysis of the full-length NP and NTD.

(A) – SDS-PAGE analysis of the NTD expression, induction with 1 mM IPTG. (B) –SDS-PAGE analysis of the NTD during the purification with the RNase A treatment. Target protein position is marked with the red arrow. (C), (D) – ESI-MS mass spectra of the full-length NP and NTD. (E), (F) – Size-exclusion chromatography traces of the pure NP and NP-RNA, absorbance detection at 280 nm (red) and 260 nm (blue). Mobile phase – 500 mM NaCl, 20 mM sodium phosphate, 100 mM imidazole. The protein molecular mass in the main peak was determined by the MALS detector (left value) and calculated from the calibration curve (right value), the molecular mass of the right peak was calculated from the calibration curve.
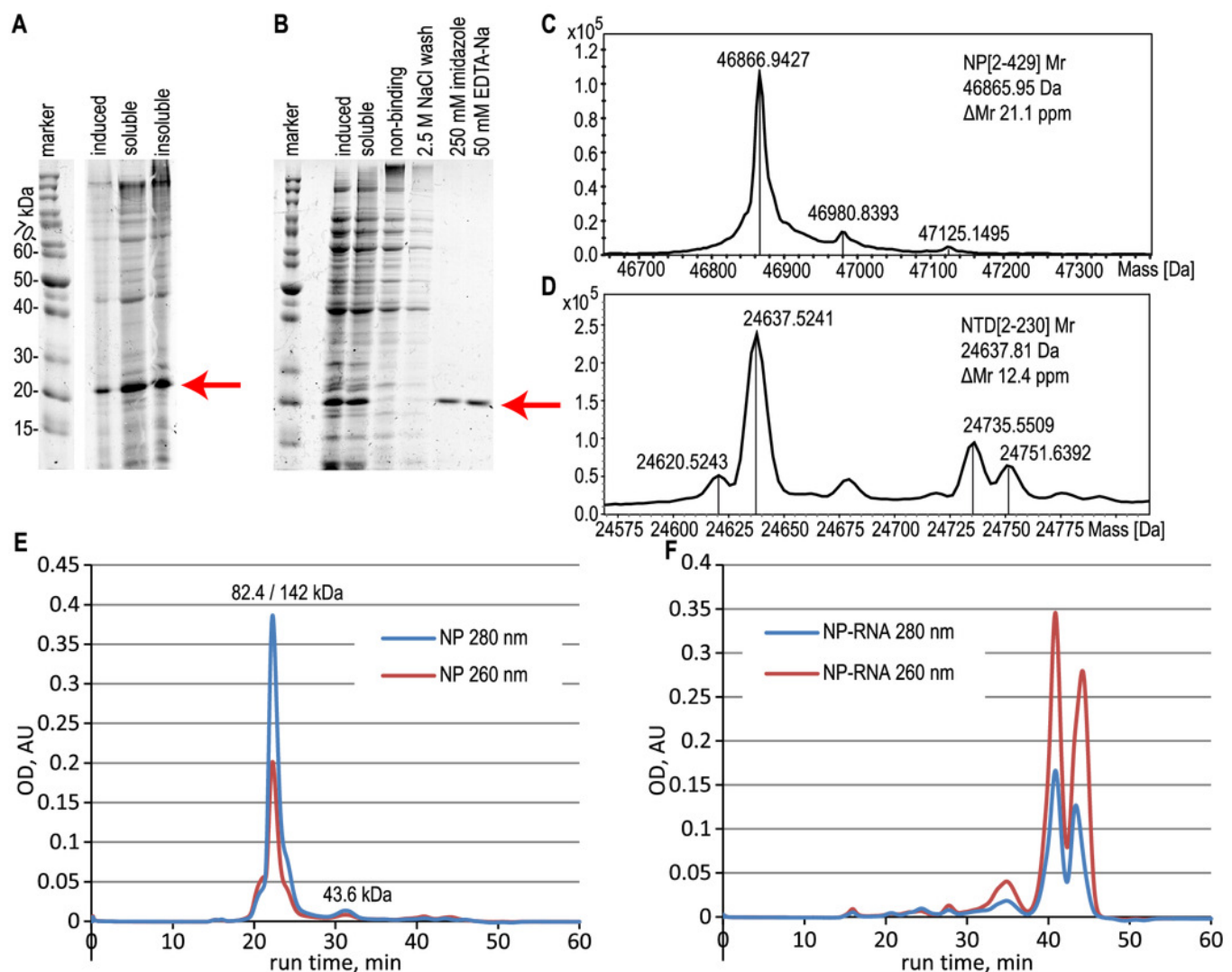
Figure 3

Figure 3 . Antigenic properties of the NP and NTD proteins.

(A) – Antibody capture ELISA with the pure NP and NP-RNA antigens, pooled PCR+ sera. (B) – antibody capture ELISA with the pure NTD and NTD-RNA antigens, pooled PCR+ sera. (C) - antibody capture ELISA with the pure NP and NP-RNA antigens, pooled pre-COVID sera. (D) – Positivity indices of the control serum samples from the NIBSC sera panel, samples dilution 1600x; NEG – negative samples; POS – positive samples. (E) – same to (D), samples dilution 400x. Statistical analysis by the paired t-test for all panels, n=3, * - p < 0.05; ** - p < 0.01; *** - p < 0.001.
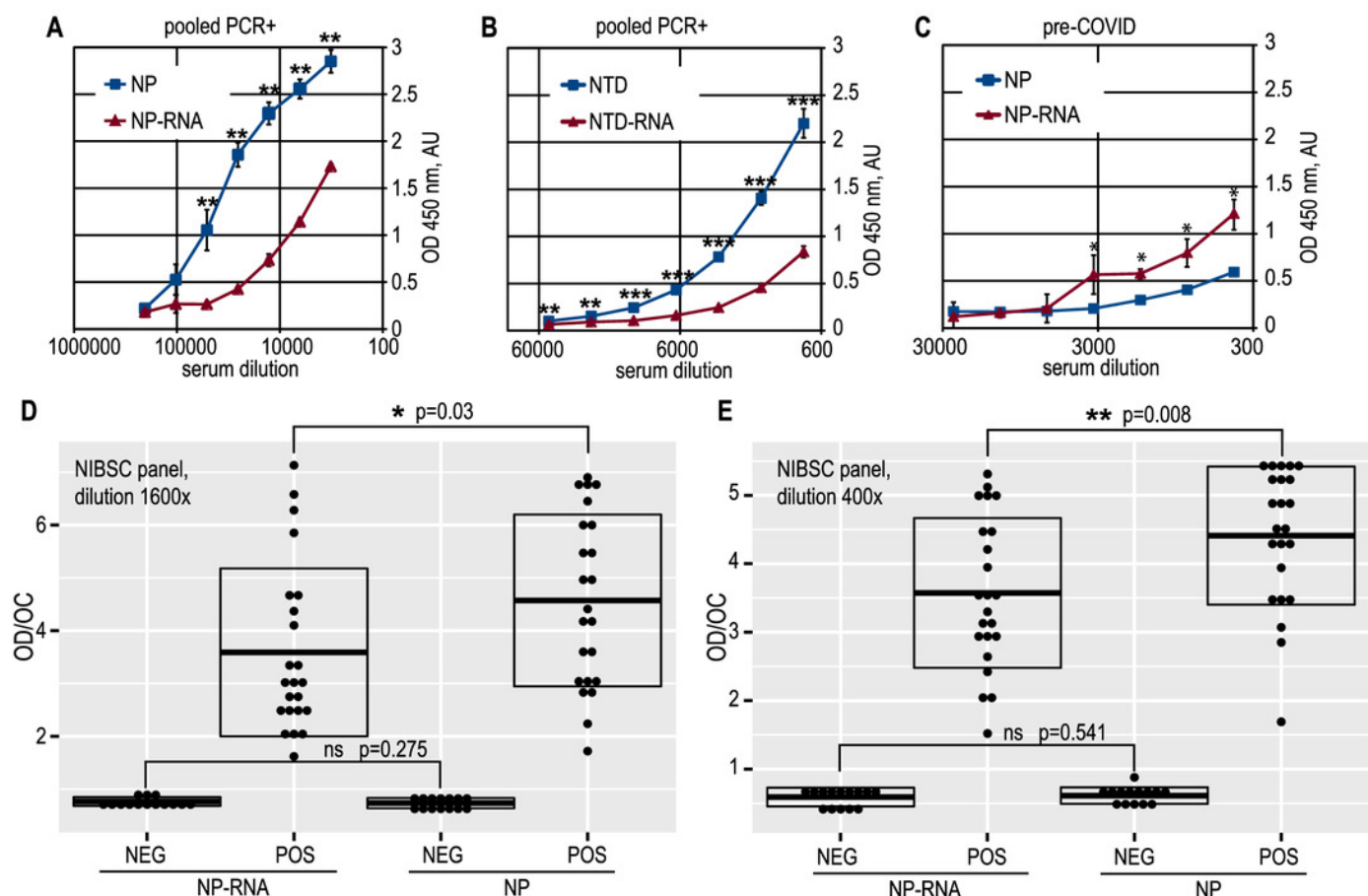
Figure 4

Correlation analysis of the NP ELISA and reference assays for NIBSC control sera samples.

The plot shows the Pearson coefficient that was used for the estimation, as well as the linear approximation. Pure NP and NP-RNA – experimentally performed assays, other data were taken from the NIBSC control sera panel description. S1, RBD, S1/S2 NP #1, NP #2 assay codes are described in detail in the Supporting Table 1. (A) – correlation analysis for the positive sample's subset; (B) – correlation analysis for all samples.

A

NIBSC panel, positive samples

B

NIBSC panel, all samples