Identification of immunodominant epitopes on nucleocapsid and spike proteins of the SARS-CoV-2 in Iranian COVID-19 patients

Faezeh Maghsood[1], Mohammad-Reza Shokri[1], Mahmood Jeddi-Tehrani[2], Monireh Torabi Rahvar[3], Abbas Ghaderi[4], Vahid Salimi[5], Gholam Ali Kardar[6], Amir-Hassan Zarnani[1], Mohammad Mehdi Amiri*[1], Fazel Shokri*[1]

1. *Department of Immunology, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran*
2. *Monoclonal Antibody Research Center, Avicenna Research Institute, ACECR, Tehran, Iran*
3. *Department of Immunology, School of Medicine, Tehran University of Medical Sciences, Tehran, Iran.*
4. *Cancer Research Center, Shiraz University of Medical Sciences, Shiraz, Iran.*
5. *Department of Virology, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran*
6. *Immunology Asthma & Allergy Research Institute, Children's Medical Center, Tehran University of Medical Sciences, Tehran, Iran*

Corresponding authors:

Professor Fazel Shokri and Dr. Mohammad Mehdi Amiri, Department of Immunology, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran; P O Box:1417613151, Tel/Fax: (+98) 2188953021; Email: fshokri@tums.ac.ir, m_amiri@tums.ac.ir

Running title: Pepscan analysis of SARS-CoV-2 RBD and NP proteins

Abstract

Given the emergence of SARS-CoV-2 virus as a life-threatening pandemic, identification of immunodominant epitopes of the viral structural proteins, particularly the nucleocapsid (NP) protein and receptor binding domain (RBD) of spike protein, is important to determine targets for immunotherapy and diagnosis. In this study, epitope screening was performed using a panel of overlapping peptides spanning the entire sequences of the RBD and NP proteins of SARS-CoV-2 in the sera from 66 COVID-19 patients and 23 healthy subjects by enzyme-linked immunosorbent assay (ELISA). Our results showed that while reactivity of patients' sera with reduced recombinant RBD protein was significantly lower than the native form of RBD ($p<0.001$), no significant differences were observed for reactivity of patients' sera with reduced and non-reduced NP protein. Pepscan analysis revealed weak to moderate reactivity towards different RBD peptide pools, which was more focused on peptides encompassing aa 181-223 of RBD. NP peptides, however, displayed strong reactivity with a single peptide covering aa 151-170. These findings were confirmed by peptide depletion experiments using both ELISA and Western blotting. Altogether, our data suggest involvement of mostly conformational disulfide bond-dependent immunodominant epitopes in RBD-specific antibody response, while the IgG response to NP is dominated by linear epitopes. Identification of dominant immunogenic epitopes in NP and RBD of SARS-CoV-2 could provide important information for the development of passive and active immunotherapy as well as diagnostic tools for the control of COVID-19 infection.

Keywords: COVID-19; Nucleocapsid; Pepscan; Receptor binding domain; SARS-CoV-2

1. Introduction

COVID-19 has remained a major health concern since the World Health Organization (WHO) declared it a pandemic on 11 March 2020 (Cohen & Normile, 2020). As of 16 May 2021, severe acute respiratory syndrome (SARS) coronavirus 2 (SARS-CoV-2) has infected more than 161 million people worldwide, with over 3 million deaths (World-Health-Organization, 2021). This has imposed a huge burden on health care systems worldwide and provoked a serious economic crisis following inevitable social lockdown (McKee & Stuckler, 2020). Thus, there is a desperate need for developing sensitive and robust diagnostic tests and potentially effective vaccines for SARS-CoV-2 to alleviate the growing socioeconomic impacts. To this goal, obtaining illuminating insights into the immune response, particularly the antibody response, to SARS-CoV2 is obviously crucial.

From the experience with SARS-CoV virus, it is postulated that neutralizing antibodies elicited against SARS-CoV-2 in convalescent individuals may be important correlates of protection (induced either by natural infection or vaccination) (Chen *et al.*, 2020, Shen *et al.*, 2020), and detection of virus-specific antibodies may provide a robust diagnostic tool for epidemiological purposes. Moreover, early seroconversion and high antibody titers could be used as a prognostic marker for disease severity and ARDS development in patients with SARS-CoV-2 (Shen *et al.*, 2020).

Like other coronaviruses infecting humans, SARS-CoV-2 is an enveloped positive-sense RNA virus, composed of four structural proteins known as spike (S), envelope (E), membrane (M), and nucleocapsid (NP) proteins (Huang *et al.*, 2020). Apart from NP which is associated to the viral RNA genome, the other three, S, E, and M proteins organize the viral envelope and are directly accessible to the host immune system (Huang *et al.*, 2020). The S protein is a large type I transmembrane glycoprotein that is responsible for receptor binding and membrane fusion. With a total length of 1273 amino acids, S protein has two functional domains, S1 and S2, located in the N- and C-terminal, respectively (Walls *et al.*, 2020, Wrapp *et al.*, 2020). S1 subunit itself consists of an N-terminal domain (14–305 residues) and a receptor-binding domain (RBD, 319–541 residues) which binds to the host receptor angiotensin converting enzyme 2 (ACE2), while the S2 subunit which plays an important role in SARS-CoV-2 fusion with the target cells, harbors the fusion peptide (FP) (788–806 residues), heptapeptide repeat sequence 1 (HR1) (912–984 residues), HR2 (1163–1213 residues), transmembrane (TM) domain (1213–1237 residues), and cytoplasmic domain (1237–1273 residues) (Chen *et al.*, 2020, Walls *et al.*, 2020).

It has become evident that SARS-CoV-2 like other human-infecting coronaviruses (e.g. SARS-CoV, MERS-CoV) can elicit IgM and IgG antibodies directed against the S and NP proteins. after initial infection (Ni *et al.*, 2020). This antibody response is initially used in serological assays for diagnostic purposes (Krammer & Simon, 2020). However,

due to the presence of several cross-reactive epitopes in NP and S with SARS-CoV and MERS-CoV (Grifoni *et al.*, 2020), the detected antibody response may not represent only anti-SARS-CoV-2 antibodies and could be a reflection of previous exposures to other human coronaviruses (Lv *et al.*, 2020). Moreover, these serology-based assays do not offer any clue about antibody functionality, including neutralization potential which is a prerequisite in vaccine development. On the other hand, antibody-dependent enhancement (ADE) through low affinity, low quantity, and non-neutralizing antibodies is a concern with human coronaviruses (Iwasaki & Yang, 2020), which further necessitates epitope dissection in order to resolve virus-specific and favorable antibodies from unfavorable cross-reactive antibodies.

Since SARS-CoV-2 virus largely depends on the spike glycoprotein for binding to ACE2 receptor and cell entry (Huang *et al.*, 2020), several studies have focused on detecting antibodies against S protein and its RBD to assess their neutralizing effect. However, antibodies against nucleoprotein appear earlier in the course of seroconversion (Meyer *et al.*, 2014) and could be more indicative for diagnostic applications. In the current study, we present several immunodominant epitopes on the RBD and NP proteins by Pepscan analysis on the sera from Iranian SARS-CoV-2 patients. This information may be useful for designing COVID-19 diagnostics and vaccines.

2. Materials and Methods

2.1. Patients and healthy controls

For primary screening, the total number of 120 cases, including 97 COVID-19 patients and 23 healthy controls, were recruited in this study. Healthy control samples were collected from the Blood Transfusion Organization of Iran 5 months before the COVID-19 epidemic in Iran and stored at -20 °C. All patients were laboratory-confirmed positive for SARS-CoV-2 by real time PCR using throat or nasopharyngeal swab specimens. All the samples were collected in April and May 2020 before the emergence of new variants of concern or interest. At that time, the original Chinese isolate (Wuhan-Hu-1) was circulating in Iran and other countries in the Middle East. Patients' sera were collected during routine clinical testing and incubated at 56 °C for 30 minutes to inactivate viruses and then stored at -20 °C until use. The mean age and standard deviation of patients and controls were 58.6±15.8 and 50.5±5.1, respectively. The patients group included 63 male and 34 female subjects, and the healthy control group included 18 males and 5 female individuals. All patients were admitted to Imam Khomeini hospital affiliated to Tehran University of Medical Sciences. Physical examination and pulse oximetry were performed for all patients, and chest CT scanning was performed for suspicious cases. Patients with the following criteria were admitted; 1) saturation of oxygen (SpO2) level <93% without supplementary oxygen and 2)

characteristics of COVID-19 disease on the patient's CT scan, i.e., involvement of more than half of the lungs. Also, patients with comorbidities and severe symptoms or signs of COVID-19 on their CT scan with normal SpO2 levels were admitted. Major clinical manifestations observed in patients include fever, cough, dyspnea, and low blood oxygen saturation levels. Serum samples were collected from all patients 3-7 days after hospital admission. The interval between the onset of symptoms and hospital admission was 11 ± 4 days.

Patients were classified into severe and non-severe groups. Severe illness included patients who had SpO2 <60% or required intensive care unit (ICU) or received invasive (mechanical) ventilation. Altogether, 49 patients (50.5%) were classified as severe, and 48 patients (49.5%) were non-severe.

Subsequently, 66 of 97 screened patients sera with the highest levels of NP and RBD-specific antibodies were selected for Pepscan analysis. The mean age and standard deviation of the selected patients were 59.6±15.1. They included 44 male and 22 female patients which 45 (68.2%) were in severe condition of disease and 21 (31.8%) were non-severe.

This study was approved by the Ethical Committee of the National Institute for Medical Research Development (NIMAD) of Iran (IR.NIMAD.REC.1399.194). Written consent was obtained from all patients included in this study or their legal representatives.

2.2. Linear peptide synthesis

The sequences we used for the design of linear peptides of the RBD and NP of SARS-CoV-2 (starin Wuhan-Hu-1) were obtained from GenBank under accession numbers NC_045512.2 and MN908947.3. For epitope screening of RBD and NP, we designed panels of linear peptides spanning the entire sequence of the RBD and NP protein of SARS-CoV-2 (each peptide contains 20 amino acid residues with 5 residues overlapping with the adjacent peptides). All the peptides were synthesized by Pepmic company (China). Peptide synthesis was performed using a standard solid-phase Fmoc method. Peptides were purified to homogeneity (purity more than 90%) by HPLC and identified by laser desorption mass spectrometry. Peptides were used individually or as pooled sets. Three peptides were combined to form one pooled peptide set. To obtain a stock solution, lyophilized individual peptides were dissolved in deionized water or Dimethyl Sulfoxide 3% (DMSO, Sigma Aldrich, Germany) according to the manufacturer's instruction. Amino acid sequences of peptide sets are shown in table 1.

2.3. Detection of IgM and IgG anti-RBD and -NP antibodies by enzyme-linked immunosorbent assay (ELISA)

For ELISA assays, we used recombinant SARS-CoV-2 NP protein expressed in Baculovirus-insect cells (Cat# 40588-V08B, Sino Biological Inc, China), and RBD expressed in HEK293 cells (Cat# 40592-V08H, Sino Biological Inc). RBD (1.0 µg/ml) or NP protein (1.5 µg/ml) were coated in flat-bottom 96-well Maxisorp microtiter plates (Nunc, Denmark) in phosphate-buffered saline (PBS) (pH 7.4) overnight at 4°C. After blocking with 3% w/v skimmed milk (Sigma Aldrich) in PBST (0.05% v/v Tween-20 in PBS, Sigma-Aldrich) as the blocking buffer, 1:200 diluted sera were applied onto the plates in the blocking buffer and incubated at 37°C for 1 h, followed by three washes with PBST. Subsequently, plates were incubated with horse-radish peroxidase (HRP)-conjugated mouse monoclonal anti-human IgM or IgG (produced in our lab) at 1:1000 and 1:2000 dilution, respectively in the blocking buffer at 37°C for 1 h. The reactions were developed by the addition of tetramethylbenzidine (TMB) substrate solution (Pishtaz Teb, Iran) for 15 minutes and stopped by the addition of 1M H_2SO_4 . Then, the optical density (OD) of the reactions was measured at 450 nm wavelength minus 630 nm using a microplate reader (Biotek, USA). The quantitative cut-off value for seropositivity for COVID-19 was defined as the mean OD of healthy samples plus 2 SDs.

2.4. Peptide-based ELISA

Preliminary peptide-based ELISA was performed using pooled sets of three peptides. An ELISA assay was initially performed using serial dilutions of a number of samples to exclude background signals and determine a dilution in which negative and positive samples could be differentiated. Based on these preliminary results, we selected a single 1:200 dilution for all samples to be able to compare the results. ELISA was performed according to the protocol described in section 2.3. with the following modifications. In brief, flat-bottom 96-well Maxisorp plates (Nunc) were coated with 3.5 µg/ml of each peptide in 3 peptide pool sets, native proteins (1.0 µg/ml of RBD or 1.5 µg/ml NP protein) or 1% 2-Mercaptoethanol (2-ME, Sigma Aldrich) reduced proteins in PBS and incubated overnight at 4 °C. After washing three times with 0.05% PBST, the plates were blocked using blocking buffer for 1 h at 37 °C, before the addition of serum samples. in the blocking buffer for 1 h at 37°C. HRP-conjugated mouse monoclonal anti-human IgG antibody was used at 1:2000 dilution for detection of peptide-bound antibodies. Development was performed by the addition of TMB and stopped with 1M H_2SO_4.

2.5. Peptide adsorption of patients' sera

To verify reactivity of RBD-specific and NP-specific antibodies to the identified peptides from both proteins, a depletion assay was performed. Using principles similar to the previous section, we performed peptide adsorption as follows. P151-170 of NP and P204-223 of RBD were selected as the immunodominant peptides. P136-155 of NP and P196-215 of RBD were selected as non-reactive control peptides. 2.0 µg/ml of selected peptides were incubated with the patient's sera for 3 h at 37 °C to adsorb peptide-specific antibodies. Peptide-adsorbed patients' sera were then incubated for 1h at 37°C in NP or RBD pre-coated ELISA plates. ELISA analysis was performed as described in section 2.4. Non-adsorbed sera were added as control.

2.6. Western blotting

To analyze reactivity of peptide-adsorbed sera with SARS-CoV-2 RBD and NP proteins, 2 µg of each protein in native or 1% 2-ME reduced form were subjected to10%. Polyacrylamide gel in SDS sample buffer. Proteins were separated by electrophoresis at 100 V for 1h, then transferred to a 0.45 µm hydrophilic polyvinylidene fluoride (PVDF) membrane at 110 V for 1.5 hours. Successful transfer and equal protein loading were confirmed by Ponceau S (Sigma Aldrich) (0.1% Ponceau S in 5% acetic acid) and Coomassie Brilliant Blue (Sigma Aldrich) staining (0.1% Brilliant blue in 50% methanol, 10% acetic acid, 40% dH_2O) on PVDF membrane and polyacrylamide gel, respectively. Membranes were blocked in 5% skimmed milk in PBS overnight at 4°C, after de-staining with H_2O. Subsequently, the membranes were incubated with serum samples diluted in blocking solution for 45 minutes at room temperature (RT). After washing for 5 times with PBST, membranes were incubated with secondary HRP-conjugated mouse monoclonal anti-human IgG at RT for 45 minutes, followed by washing 5 times for 5 minutes. Finally, positive signals were detected by chemiluminescence ECL Prime (GE Healthcare, Sweden) solution.

2.7. Statistical analysis

GraphPad Prism software, version 8 (GraphPad Software Inc., La Jolla, USA) was used for plotting graphs and statistical analysis. All values were expressed as mean ± standard deviation (SD). Statistical analysis was performed using Mann–Whitney U test for Pepscan analysis and Wilcoxon test for comparing reactivity against reduced and native target

proteins and depletion assays analysis. Differences were considered to be significant at P<0.05(*), P<0.01(**), P<0.001(***), and P<0.0001(****).

Results

3.1. IgM and IgG responses to RBD and NP proteins

Serum levels of IgM and IgG against RBD and NP proteins were determined in patients' sera (n=97) by ELISA (Fig. 1). Sera from 23 healthy individuals were used as controls. Our data demonstrated that based on assigned cut-off OD values (cut-off =mean ±2SD of normal individuals) anti-RBD and anti-NP IgG were positive in 94% and 92%. and anti-RBD and anti-NP IgM were positive in 90% and 80% of patients' sera, respectively. Subsequently, 66 serum samples from patients with IgG titer higher than the cut-off level (0.33 for anti-RBD IgG and 0.43 for anti-NP IgG) were selected for further Pepscan experiments.

2.2. Identification of linear epitopes on the SARS-CoV-2 RBD and NP proteins

We next assessed the linear dominant antigenic determinants within the RBD domain of the spike protein (319-541 aa), and NP protein (1-419 aa) by Pepscan analysis. A total of 66 serum samples from COVID-19 patients were selected for epitope mapping which had been verified to be reactive to the target proteins by ELISA. Serum samples from 23 healthy subjects were also included as negative controls. A set of 20-aa overlapping peptides spanning the entire RBD domain and NP protein in pools of 3 adjacent peptides were used as coating antigens in ELISA (Table 1), and serum samples were tested against each of the peptides. To test whether the reactivity is proportional to linear epitopes or conformational ones, reduced as well as native proteins were also used as coating antigens, and all sera were assessed with these antigens as well.

2.2.1. *Reactivity against native and reduced SARS-CoV-2 RBD proteins*

Native and reduced SARS-CoV-2 RBD proteins were employed to assess the reactivity of RBD-specific IgG antibodies with cysteine bonds dependent and independent epitopes. Significantly higher reactivity was observed with native non-reduced protein as compared to the reduced preparation (*p=0.0002*) (Fig. 2).

2.2.2. ***Reactivity against RBD peptide pools***

While all 66 serum samples from COVID-19 patients were reactive against native RBD protein, less than 40% of these samples recognized peptide pools A to D, and pool E alone reacted with about 40% of patients' sera. On the other hand, reactivity with all RBD peptide pools was significantly higher in patients' sera than healthy controls (Fig. 3a).

2.2.3. ***Identification of the immunodominant peptide in pool E***

Since patients' sera showed better reactivity with pool "E" of RBD peptides, we evaluated the reactivity of 25 pool "E" positive serum samples with individual peptides of this peptide pool. Most samples reacted either with peptide P204-223 or 181-200 of RBD (Fig. 3b).

2.2.4. ***Depletion assay confirmed specific reactivity with P204-223 on SARS-CoV-2 RBD protein***

Further assessment was performed to verify specific reactivity of RBD-specific antibodies from 10 patients to P204-223 peptide from RBD protein (Fig. 4). Patients' sera were adsorbed with P204-223 as the most reactive peptide and P196-215 as the non-reactive peptide, and evaluated by ELISA against native RBD (Fig. 4a). Two of these samples were also tested by Western blotting (Fig. 4b). The reactivity of non-adsorbed sera against RBD was slightly, but non-significantly, higher in comparison with that of P204-223 adsorbed sera ($p=0.274$).

2.3. Identification of linear epitopes on the SARS-CoV-2 NP protein

2.3.1. ***Reactivity against native and reduced SARS-CoV-2 NP proteins***

Native and reduced SARS-CoV-2 NP proteins were employed to assess the reactivity of NP-specific antibodies with cysteine bonds dependent and independent epitopes. The reactivity to native NP was similar to the reduced NP ($p=0.206$) (Fig. 5), suggesting that denaturation does not significantly alter recognition of the NP protein by patients' antibodies.

2.3.2. ***Reactivity against NP peptide pools***

Pepscan analysis of peptide pools covering the entire NP protein against sera obtained from COVID-19 and healthy controls also revealed significantly higher reactivity of patients' sera than healthy controls in each block of peptides (Fig. 6a). Most pools of peptides were recognized by less than 50% of patients' sera, while one distinct antigenic site corresponding to aa 136-185 in the N-terminal domain of NP protein (pool I) reacted very strongly with more than 75% of the serum samples from COVID-19 patients. This suggests that the identified region, namely pool I, contains at least one of the major linear immunodominant epitopes that induces the antibody response in COVID-19 patients. The results also indicate that the reactivity against pool I was collectively weaker than reactivity to native NP protein (Fig. 6a), which may reflect the contribution of other potential conformational epitopes in native NP.

2.3.3. ***Identification of the immunodominant peptide in pool I***

Next, we further assessed individual peptides within pool I from NP protein which includes aa 136-155, 151-170, and 166-185, to precisely determine the exact 20-mer epitope which attributes to the highest reactivity of COVID-19 patients' sera (Fig. 6b). Sera obtained from 25 patients were used in this experiment. The data revealed that peptide P151-170 is the dominant hit from pool "I".

2.3.4. ***Depletion assay confirmed specific reactivity with P151-170 on SARS-CoV-2 NP protein***

Further assessment was performed to verify the specific reactivity of NP-specific antibodies from 10 patients to the dominant P151-170 peptide from NP protein (Fig. 7). Patients' sera were individually adsorbed with P151-170, as well as P136-155 as the non-reactive peptide, and evaluated by ELISA against native NP (Fig. 7a). Two of these samples were also tested by Western blotting (Fig. 7b). The data revealed that sera adsorbed with peptide P151-170 have lower reactivity ($p=0.0156$) with native NP in comparison with non-adsorbed sera.

3. Discussion

Identification of major antigenic determinants of SARS-CoV-2 proteins which provoke remarkable antibody response in COVID-19 patients may provide valuable information for understanding the virus-neutralizing antibody response and developing efficient vaccines and serological assays.

Here, we have mapped the linear immunodominant epitopes on SARS-CoV-2 NP and RBD proteins using serum samples collected from COVID-19 Iranian patients by Pepscan analysis employing 20-aa overlapping peptides spanning the whole sequence of both proteins. We showed that Iranian patients with COVID-19 develop significant anti-RBD and anti-NP IgG and IgM antibody responses. However, while the antibody response against RBD seems to be largely raised against the S-S bond-dependent conformational determinants, the antibody response against NP protein is mostly directed to linear epitopes. Nevertheless, the limited size of samples and lack of information about the kinetics of antibody response in patients' sera are limitations which may influence interpretetion of our results.

Despite strong serological reactivity of convalescent patients against RBD reported in studies using full-length RBD antigen of SARS-CoV-2 (Amanat et al., 2020, Zeng et al., 2021), reports on serological reactivity to "linear" immunodominant sites on SARS-CoV-2 RBD are very limited in the literature. Although Zhang et al. reported four linear immunodominant sites on RBD detected by sera from COVD-19 patients (Zhang *et al.*, 2020), we observed low reactivity to peptides covering the RBD of SARS-CoV-2. Weak serological reactivity to peptides within the RBD was also observed in a recent study using a highly multiplexed peptide assay (PepSeq) by Ladner et al. (Ladner *et al.*, 2021). Meng Poh and colleagues using a pepscan analysis, also failed to find a linear immunodominant epitope exactly localized to RBD region with high reactivity to patients' sera (Poh *et al.*, 2020). This weak/ lack of reactivity in peptide-based approaches implies that antibodies reactive to RBD region are largely directed against conformational epitopes and/or epitopes generated by post-translational modifications. Hastie et al. assessed the competitive relationship between 186 RBD-directed mAbs by high-throughput surface plasmon resonance (HT-SPR). They sorted RBD-directed antibodies into seven core “communities”. Further studies were performed by negative-stain electron microscopy (NS-EM) to understand the position of each community. Footprint mapping divided RBD epitopes into seven communities recognized by MAbs containing RBM (receptor binding motif), the outer face of the RBD, or the inner face of the RBD which were further divided into smaller clusters and bins. RBM was the largest community and the mAbs binding to the RBM potently competed with ACE2 and required the “up” position of RBD for binding (Hastie *et al.*, 2021). Based on their results indicating the dominancy of RBM-directed neutralizing antibodies, and since Pepscan analysis in our assay and peptide-based antibody assays exclusively detects linear epitopes created by the primary structure of proteins and are

blind to epitopes generated in the secondary or tertiary structure of proteins, we speculate that RBM possesses conformational epitopes, so that we found no reactivity in patient's sera against linear peptides corresponding to RBM region (P106-125, P121-140, P136-155, P151-170, P166-185, and P181-200). Interestingly, the reduction of S-S bonds within the native RBD molecule by a reducing agent resulted in a significant reduction of antibody reactivity in patients' sera (Fig. 2).

To assess the linear epitopes which provoke anti-RBD antibody response, we performed Pepscan analysis using 20-aa long overlapping peptide covering the entire RBD domain. Although antibody reactivity against 5 different peptide pools, each consisting of 3 peptides, was significantly higher in sera from patients compared to healthy subjects, it was significantly lower when compared with reactivity to the native RBD protein (Fig. 3a). Of the 5 peptide pools, pool "E" which covers the C-terminal residues of RBD (aa 181-223) displayed the highest reactivity. Thus, we investigated the antibody response to the three peptides of this pool in 25 selected patients' sera to identify the most immunodominant one. Both peptides P181-200 and particularly P204-223, but not peptide P196-215, displayed modest reactivity with some of the samples (Fig. 3b). Preincubation of serum samples from 10 of these patients with peptide P204-223 resulted in slightly lower reactivity of these samples to native RBD by ELISA and Western blot (Fig. 4). All these findings suggest that the antibody response to RBD is dominated by S-S bond-dependent conformational epitopes. The fact that the immunogenicity of RBD mainly relies on the conformational structures and/or post-translational modifications (i.e., glycosylation) is proved by studying the binding footprint of neutralizing monoclonal antibodies that inhibit RBD binding to ACE2. Studies revealed that some residues that are distal in the linear sequence of RBD and their glycosylation state contribute to antibody binding (Pinto *et al.*, 2020, Yuan *et al.*, 2020).

We adopted a similar approach to investigate the antibody response against the NP protein. SARS-CoV-2 NP is a 419 aa phosphoprotein that associates with the viral RNA genome as well as other proteins to form the ribonucleoprotein core (Mu *et al.*, 2020). Like SARS-CoV, NP protein of SARS-CoV-2 consists of three distinct domains: an N-terminal RNA-binding domain (NTD) which associates with the RNA genome, an intrinsically disordered central Ser/Arg (SR)-rich linker and a C-terminal domain (CTD) which allows dimerization of NP proteins (Kang *et al.*, 2020, Khan *et al.*, 2020).

Our results revealed that in contrast to RBD, NP-specific antibody response is mainly directed against linear epitopes. No significant difference was observed between serum levels of antibody against reduced and non-reduced NP protein (Fig. 5). This was also supported by the Pepscan data which showed substantially higher reactivity of the anti-NP antibodies from patients' sera with all peptide pools, particularly peptide pool "I", which covers aa 136-185 (Fig. 6a). When the three peptides within this pool were dissected and tested in a number of patients' sera, the antibody response was almost entirely directed against one of these three peptides which encompasses aa 151-170 (Fig. 6b).

Amino acids 151-170 are located on RNA-binding terminal domain of NP protein which has been previously reported by Amruna et al. as an immunodominant epitope in SARS-CoV-2 and as a disease severity correlate in COVID-19 patients (Amrun *et al.*, 2020). Sequence homology analysis demonstrated that the epitope we identified in SARS-CoV-2 NP protein is highly conserved among human coronavirus strains, including SARS-CoV, MERS-CoV, HKU1-CoV, and OC43-CoV. Although most studies only address the diagnostic value of nucleocapsid protein and its epitopes for detection of seroconversion in COVID-19 patients, the highly conserved nature of this epitope in NP protein would be an advantage in developing epitope-based vaccines against SARS-CoV-2 which can develop cross-protection against other human coronaviruses. The majority of vaccines which are under clinical investigation against SARS-CoV-2 are developed based on viral spike protein and mostly rely on the ensuing neutralizing antibody response which blocks viral entry rather than killing infected cells. This could be largely attributed to the localization of NP protein. While the three other structural proteins of the virus, S, M, and E reside at the interface of the virus to the external environment, the NP protein is located inside the viral particles and is not accessible to the antibody during the course of infection. Therefore, although elevated anti-SARS-CoV-2 NP IgG and IgM antibody titers were observed in our study and have been reported by other investigators (Guo *et al.*, 2020, To *et al.*, 2020), they do not seem to have neutralization potential. However, anti-NP humoral responses might exert other immune effector antibody-dependent mechanisms on virus-infected cells displaying NP to eliminate the viral infection. Functional characterization of anti-SARS-CoV-2 hyperimmune sera from convalescent individuals showed effective induction of antibody-dependent cellular cytotoxicity (ADCC) and antibody-dependent cellular phagocytosis (ADCP) against NP and S proteins (Díez *et al.*, 2021).

Since most of our current knowledge about SARS-CoV-2 NP protein comes from previous studies on SARS-CoV, better evidence in this regard could be obtained from SARS-CoV literature. Several studies during the SARS outbreak clearly showed that SARS-CoV NP protein could not induce strong neutralizing antibody responses neither in human nor in animal (Buchholz *et al.*, 2004, Pang *et al.*, 2004, Liang *et al.*, 2005); however, significant cytotoxic T lymphocyte (CTL) response against NP has been reported using vector-based vaccines containing NP protein (Gao *et al.*, 2003, Buchholz *et al.*, 2004, Zhu *et al.*, 2004) . Kim and colleagues introduced an effective DNA vaccine using SARS-CoV NP protein as a target antigen fused to calreticulin (CRT) to enhance MHC class I presentation of linked antigen to CD8(+) T cells. Their results showed that the NP-based vaccine generated strong NP-specific humoral and T-cell immune responses in mice (Kim *et al.*, 2004). Recently, it has been suggested that SARS-CoV-2 NP could be considered as an advantageous vaccine target owing to its conserved nature and strong immunogenicity (Dutta *et al.*, 2020). A recently published data by Ahlén et al. (Ahlén *et al.*, 2020), also showed a DNA vaccine based on a codon-optimized SARS-CoV-2 NP gene induced high titers of anti-NP antibodies in immunized rabbits, but most interestingly, they showed that immunization of

mice with a DNA vaccine expressing the SARS-CoV-2 NP protein induced the strongest T cell response against a peptide region spanning our P151-170 peptide.

4. Conclusion

In summary, we studied the B cell linear epitopes on RBD and NP proteins of SARS-CoV-2 using serum samples from COVID-19 patients. While the antibody response to NP was mainly directed against linear epitopes as evidenced by reactivity to synthetic peptides and reduced native protein, RBD specific antibodies seem to largely recognize disulphid-bond dependent conformational epitopes. These findings propose that peptide-based vaccines may not be able to elicit virus neutralizing antibodies against SARS-CoV-2. Furthermore, profiling of the NP specific antibody response by Pepscan suggests that the immunodominant peptides could be employed for immunodiagnosis of the infection.

Authors' Contributions

F.M. and MR.Sh.; performing the assays and writing original draft, M.TR.; writing original draft, M.JT, V.S., Gh. K, and AH.Z.; reviewing and editing the manuscript, conceptualization, data analysis, clinical samples and data collection and analysis, MM.A., F.Sh.; writing original draft, review and editing, project conceptualization, data validation, project administration and supervision. All authors have read and agreed to the published version of the article.

Acknowledgements

The authors would like to thank Jalal Khoshnoodi and Mohammad Ali Judaki for their technical assistance.

Funding

This study was partially supported by TUMS and a grant from the National Institute for Medical Research Development (NIMAD) of Iran (Grant No. 993421).

Ethics approval

This study was approved by the Ethical Committee of the National Institute for Medical Research Development (NIMAD) of Iran (IR.NIMAD.REC.1399.194). Written consent was obtained from all patients included in this study or their legal representatives.

Competing interests

The authors declare no competing interests.

Data availability

The original data sets of this study are available from the corresponding author, upon reasonable request.

References:

Ahlén G, Frelin L, Nikouyan N, Weber F, Höglund U, Larsson O, Westman M, Tuvesson O, Gidlund E-k & Cadossi M (2020) The SARS-CoV-2 N protein is a good component in a vaccine. *Journal of Virology* **94**.

Amanat F, Stadlbauer D, Strohmeier S, Nguyen TH, Chromikova V, McMahon M, Jiang K, Arunkumar GA, Jurczyszak D & Polanco J (2020) A serological assay to detect SARS-CoV-2 seroconversion in humans. *Nature medicine* 1-4.

Amrun SN, Lee CY-P, Lee B, Fong S-W, Young BE, Chee RS-L, Yeo NK-W, Torres-Ruesta A, Carissimo G & Poh CM (2020) Linear B-cell epitopes in the spike and nucleocapsid proteins as markers of SARS-CoV-2 exposure and disease severity. *EBioMedicine* **58**: 102911.

Buchholz UJ, Bukreyev A, Yang L, Lamirande EW, Murphy BR, Subbarao K & Collins PL (2004) Contributions of the structural proteins of severe acute respiratory syndrome coronavirus to protective immunity. *Proceedings of the National Academy of Sciences* **101**: 9804-9809.

Chen L, Xiong J, Bao L & Shi Y (2020) Convalescent plasma as a potential therapy for COVID-19. *The Lancet Infectious Diseases* **20**: 398-400.

Chen Y, Guo Y, Pan Y & Zhao ZJ (2020) Structure analysis of the receptor binding of 2019-nCoV. *Biochemical and biophysical research communications*.

Cohen J & Normile D (2020) New SARS-like virus in China triggers alarm. p.^pp. American Association for the Advancement of Science.

Díez JM, Romero C, Cruz M*, et al.* (2021) Anti-SARS-CoV-2 hyperimmune globulin demonstrates potent neutralization and antibody-dependent cellular cytotoxicity and phagocytosis through N and S proteins. *J Infect Dis*.

Dutta NK, Mazumdar K & Gordy JT (2020) The nucleocapsid protein of SARS–CoV-2: a target for vaccine development. *Journal of virology* **94**.

Gao W, Tamin A, Soloff A, D'Aiuto L, Nwanegbo E, Robbins PD, Bellini WJ, Barratt-Boyes S & Gambotto A (2003) Effects of a SARS-associated coronavirus vaccine in monkeys. *The Lancet* **362**: 1895-1896.

Grifoni A, Sidney J, Zhang Y, Scheuermann RH, Peters B & Sette A (2020) A sequence homology and bioinformatic approach can predict candidate targets for immune responses to SARS-CoV-2. *Cell host & microbe*.
Guo L, Ren L, Yang S, Xiao M, Chang D, Yang F, Dela Cruz CS, Wang Y, Wu C & Xiao Y (2020) Profiling early humoral response to diagnose novel coronavirus disease (COVID-19). *Clinical Infectious Diseases*.
Hastie KM, Li H, Bedinger D*, et al.* (2021) Defining variant-resistant epitopes targeted by SARS-CoV-2 antibodies: A global consortium study. *Science* **374**: 472-478.
Huang Y, Yang C, Xu X-f, Xu W & Liu S-w (2020) Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19. *Acta Pharmacologica Sinica* 1-9.
Iwasaki A & Yang Y (2020) The potential danger of suboptimal antibody responses in COVID-19. *Nature Reviews Immunology* 1-3.
Kang S, Yang M, Hong Z, Zhang L, Huang Z, Chen X, He S, Zhou Z, Zhou Z & Chen Q (2020) Crystal structure of SARS-CoV-2 nucleocapsid protein RNA binding domain reveals potential unique drug targeting sites. *Acta Pharmaceutica Sinica B*.
Khan MT, Zeb MT, Ahsan H, Ahmed A, Ali A, Akhtar K, Malik SI, Cui Z, Ali S & Khan AS (2020) SARS-CoV-2 nucleocapsid and Nsp3 binding: an in silico study. *Archives of Microbiology* 1-8.
Kim TW, Lee JH, Hung C-F, Peng S, Roden R, Wang M-C, Viscidi R, Tsai Y-C, He L & Chen P-J (2004) Generation and characterization of DNA vaccines targeting the nucleocapsid protein of severe acute respiratory syndrome coronavirus. *Journal of virology* **78**: 4638-4645.
Krammer F & Simon V (2020) Serology assays to manage COVID-19. *Science* **368**: 1060-1061.
Ladner JT, Henson SN, Boyle AS*, et al.* (2021) Epitope-resolved profiling of the SARS-CoV-2 antibody response identifies cross-reactivity with endemic human coronaviruses. *Cell reports Medicine* **2**: 100189.
Liang M-f, Du R-l, Liu J-z, Li C, Zhang Q-f, Han L-l, Yu J-S, Duan S-m, Wang X-f & Wu K-x (2005) SARS patients-derived human recombinant antibodies to S and M proteins efficiently neutralize SARS-coronavirus infectivity. *Biomedical and Environmental Sciences* **18**: 363.
Lv H, Wu NC, Tsang OT-Y, Yuan M, Perera RA, Leung WS, So RT, Chan JMC, Yip GK & Chik TSH (2020) Cross-reactive antibody response between SARS-CoV-2 and SARS-CoV infections. *Cell Reports* 107725.
McKee M & Stuckler D (2020) If the world fails to protect the economy, COVID-19 will damage health not just now but also in the future. *Nature Medicine* **26**: 640-642.
Meyer B, Drosten C & Müller MA (2014) Serological assays for emerging coronaviruses: challenges and pitfalls. *Virus research* **194**: 175-183.
Mu J, Xu J, Zhang L, Shu T, Wu D, Huang M, Ren Y, Li X, Geng Q & Xu Y (2020) SARS-CoV-2-encoded nucleocapsid protein acts as a viral suppressor of RNA interference in cells. *Science China Life Sciences* 1-4.
Ni L, Ye F, Cheng M-L, Feng Y, Deng Y-Q, Zhao H, Wei P, Ge J, Gou M & Li X (2020) Detection of SARS-CoV-2-specific humoral and cellular immunity in COVID-19 convalescent individuals. *Immunity*.
Pang H, Liu Y, Han X, Xu Y, Jiang F, Wu D, Kong X, Bartlam M & Rao Z (2004) Protective humoral responses to severe acute respiratory syndrome-associated coronavirus: implications for the design of an effective protein-based vaccine. *Journal of general virology* **85**: 3109-3113.
Pinto D, Park Y-J, Beltramello M, Walls AC, Tortorici MA, Bianchi S, Jaconi S, Culap K, Zatta F & De Marco A (2020) Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody. *Nature* 1-6.
Poh CM, Carissimo G, Wang B, Amrun SN, Lee CY-P, Chee RS-L, Fong S-W, Yeo NK-W, Lee W-H & Torres-Ruesta A (2020) Two linear epitopes on the SARS-CoV-2 spike protein that elicit neutralising antibodies in COVID-19 patients. *Nature Communications* **11**: 1-7.
Shen C, Wang Z, Zhao F, Yang Y, Li J, Yuan J, Wang F, Li D, Yang M & Xing L (2020) Treatment of 5 critically ill patients with COVID-19 with convalescent plasma. *Jama* **323**: 1582-1589.
To KK-W, Tsang OT-Y, Leung W-S, Tam AR, Wu T-C, Lung DC, Yip CC-Y, Cai J-P, Chan JM-C & Chik TS-H (2020) Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. *The Lancet Infectious Diseases*.
Walls AC, Park Y-J, Tortorici MA, Wall A, McGuire AT & Veesler D (2020) Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell*.
World-Health-Organization (2021) WHO Coronavirus (COVID-19) Dashboard. p.^pp.

Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh C-L, Abiona O, Graham BS & McLellan JS (2020) Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**: 1260-1263.
Yuan M, Wu NC, Zhu X, Lee C-CD, So RT, Lv H, Mok CK & Wilson IA (2020) A highly conserved cryptic epitope in the receptor binding domains of SARS-CoV-2 and SARS-CoV. *Science* **368**: 630-633.
Zeng W, Ma H, Ding C, Yang Y, Sun Y, Huang X, He W, Xiang Y, Gao Y & Jin T (2021) Characterization of SARS-CoV-2-specific antibodies in COVID-19 patients reveals highly potent neutralizing IgA. *Signal transduction and targeted therapy* **6**: 35.
Zhang BZ, Hu YF, Chen LL*, et al.* (2020) Mining of epitopes on spike protein of SARS-CoV-2 from COVID-19 patients. *Cell research* **30**: 702-704.
Zhu M-S, Pan Y, Chen H-Q, Shen Y, Wang X-C, Sun Y-J & Tao K-H (2004) Induction of SARS-nucleoprotein-specific immune response by use of DNA vaccine. *Immunology letters* **92**: 237-243.

Table 1. Amino acid sequences of peptide sets employed in Pepscan analysis

No.	aa position	Sequence
RBD Peptides		
1	P1-20	RVQPTESIVRFPNITNLCPF
2	P16-35	NLCPFGEVFNATRFASVYAW
3	P31-50	SVYAWNRKRISNCVADYSV
4	P46-65	DYSVLYNSASFSTFKCYGVS
5	P61-80	CYGVSPTKLNDLCFTNVYAD
6	P76-95	NVYADSFVIRGDEVRQIAPG
7	P91-110	QIAPGQTGKIADYNYKLPDD
8	P106-125	KLPDDFTGCVIAWNSNNLDS
9	P121-140	NNLDSKVGGNYNYLYRLFRK
10	P136-155	RLFRKSNLKPFERDISTEIY
11	P151-170	STEIYQAGSTPCNGVEGFNC
12	P166-185	EGFNCYFPLQSYGFQPTNGV
13	P181-200	PTNGVGYQPYRVVVLSFELL
14	P196-215	SFELLHAPATVCGPKKSTNL
15	P204-223	ATVCGPKKSTNLVKNKCVNF
NP Peptides		
1	P1-20	SDNGPQNQRNAPRITFGGPS
2	P16-35	FGGPSDSTGSNQNGERSGAR
3	P31-50	RSGARSKQRRPQGLPNNTAS
4	P46-65	NNTASWFTALTQHGKEDLKF
5	P61-80	EDLKFPRGQGVPINTNSSPD
6	P76-95	NSSPDDQIGYYRRATRRIRG
7	P91-110	RRIRGGDGKMKDLSPRWYFY
8	P106-125	RWYFYYLGTGPEAGLPYGAN

9	P121-140	PYGANKDGIIWVATEGALNT
10	P136-155	GALNTPKDHIGTRNANNAA
11	P151-170	ANNAAIVLQLPQGTTLPKGF
12	P166-185	LPKGFYAEGSRGGSQASSRS
13	P181-200	ASSRSSSRSRNSSRNSTPGS
14	P196-215	STPGSSRGTSPARMAGNGGD
15	P211-230	GNGGDAALALLLLDRLNQLE
16	P226-245	LNQLESKMSGKGQQQQGQTV
17	P241-260	QGQTVTKKSAAEASKKPRQK
18	P256-275	KPRQKRTATKAYNVTQAFGR
19	P271-290	QAFGRRGPEQTQGNFGDQEL
20	P286-305	GDQELIRQGTDYKHWPQIAQ
21	P301-320	PQIAQFAPSASAFFGMSRIG
22	P316-335	MSRIGMEVTPSGTWLTYTAA
23	P331-350	TYTAAIKLDDKDPNFKDQVI
24	P346-365	KDQVILLNKHIDAYKTFPPT
25	P361-380	TFPPTEPKKDKKKKADETQA
26	P376-395	DETQALPQRQKKQQTVTLLP
27	P391-410	VTLLPAADLDDFSKQLQQSM
28	P399-418	LDDFSKQLQQSMSSADSTQA

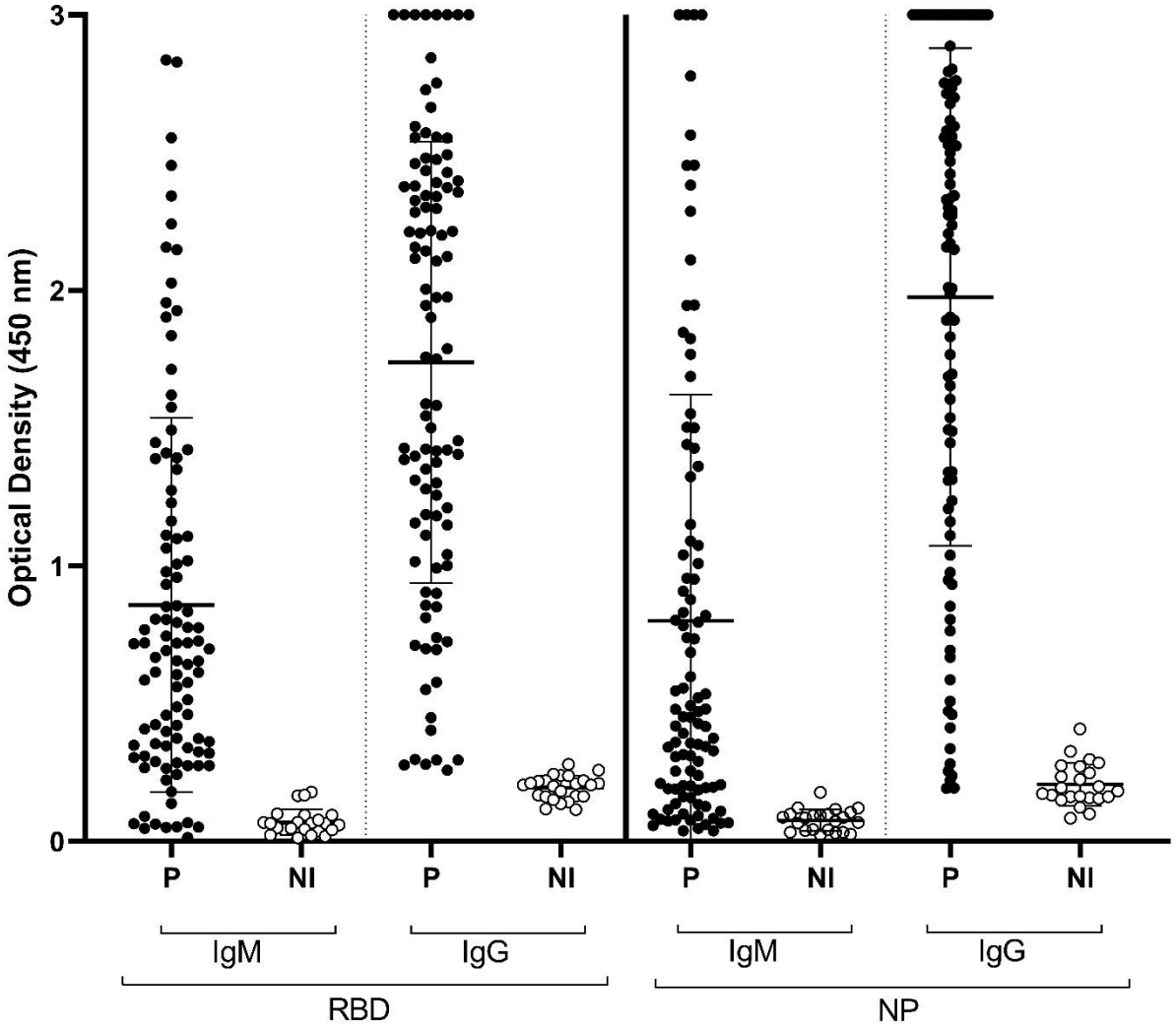

Fig 1. Detection of RBD and NP-specific IgM and IgG antibodies in serum of COVID-19 patients and normal individuals. The results represent OD values obtained by ELISA in serum from 97 patients (P) and 23 normal individuals (NI). The solid lines represent mean values ± SD. NP: Nucleocapsid protein; RBD: Receptor binding domain.

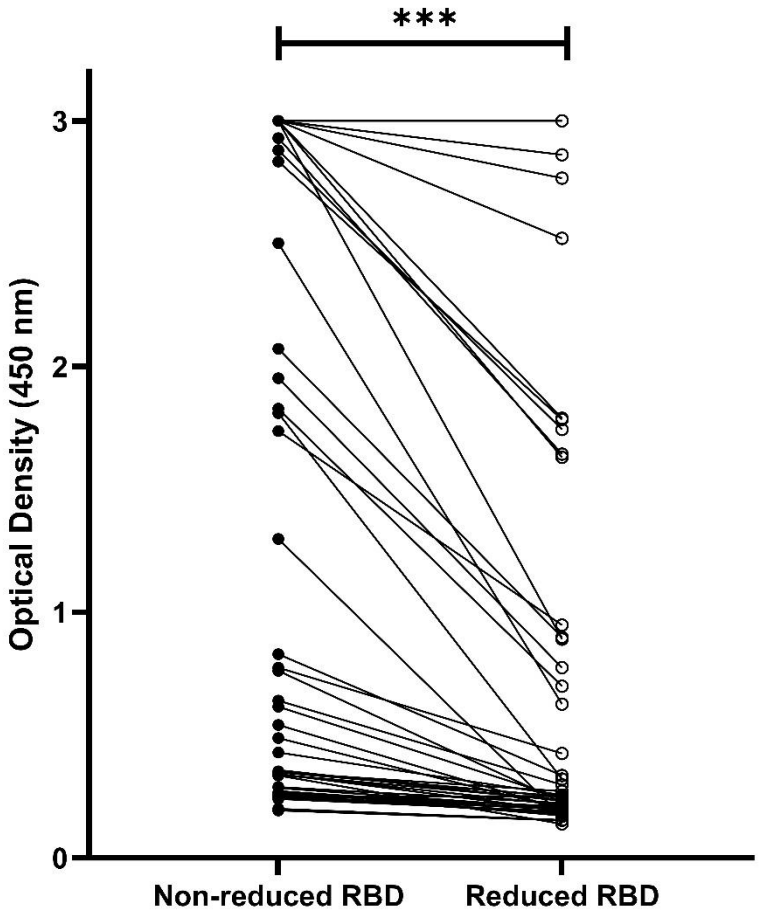

Fig 2. Reactivity of patients' sera with native and reduced SARS-CoV-2 RBD protein. Native and reduced SARS-CoV-2 RBD proteins were used for coating ELISA plates, and the sera from 48 COVID-19 patients were tested. Reactivity with native RBD was significantly stronger than the reduced protein, *p<0.001* (***). RBD: Receptor-binding domain.

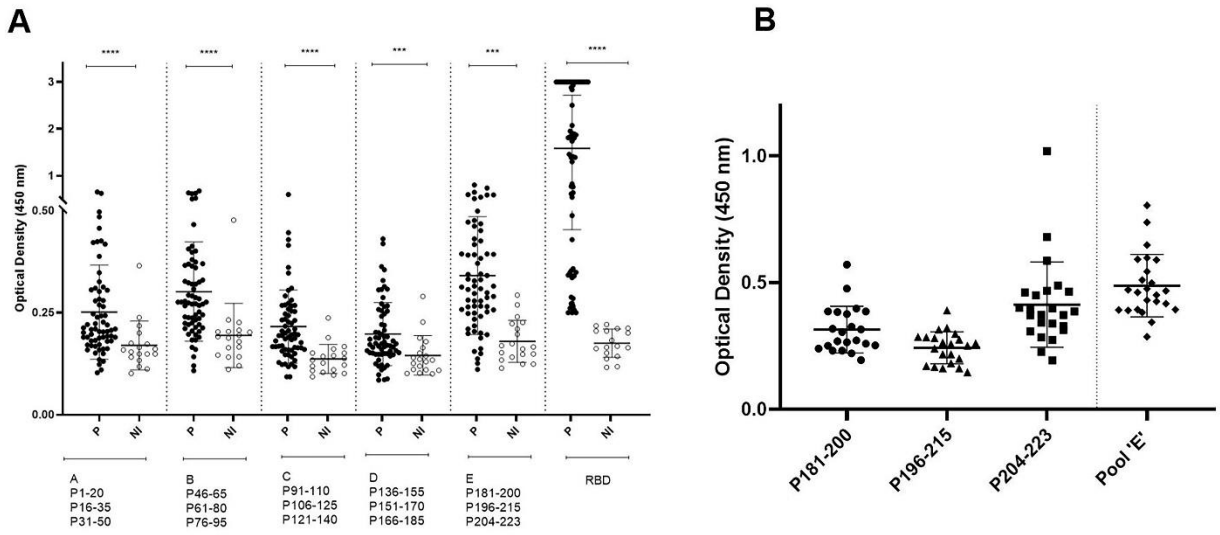

Fig 3. Mapping the immunodominant epitopes of the SARS-CoV-2 RBD protein by Pepscan analysis against antisera from COVID-19 patients and identification of the immunodominant peptide. (A) Sera of COVID-19 patients (n = 66) and healthy blood donors (n=23) were subjected to 5 pools (A to E) of overlapping peptides. Each pool containing three 20-mer adjacent peptides that covers the entire RBD domain of S protein. The reactivity of sera with the overlapping peptide pools from RBD protein was determined by ELISA**. (B)** Identification of the immunodominant epitope among "pool E" peptides from RBD protein. Individual peptides (P181-200, P196-215, and P204-223) of the positive peptide pool of RBD (pool E) were screened with plasma samples from 25 COVID-19 patients having the highest reactivity with this pool to determine the most reactive peptide. The solid lines represent mean values ± SD, *P<0.001*(***), *P<0.0001*(****). RBD: Receptor binding domain; P: Patients; NI: Normal individuals.

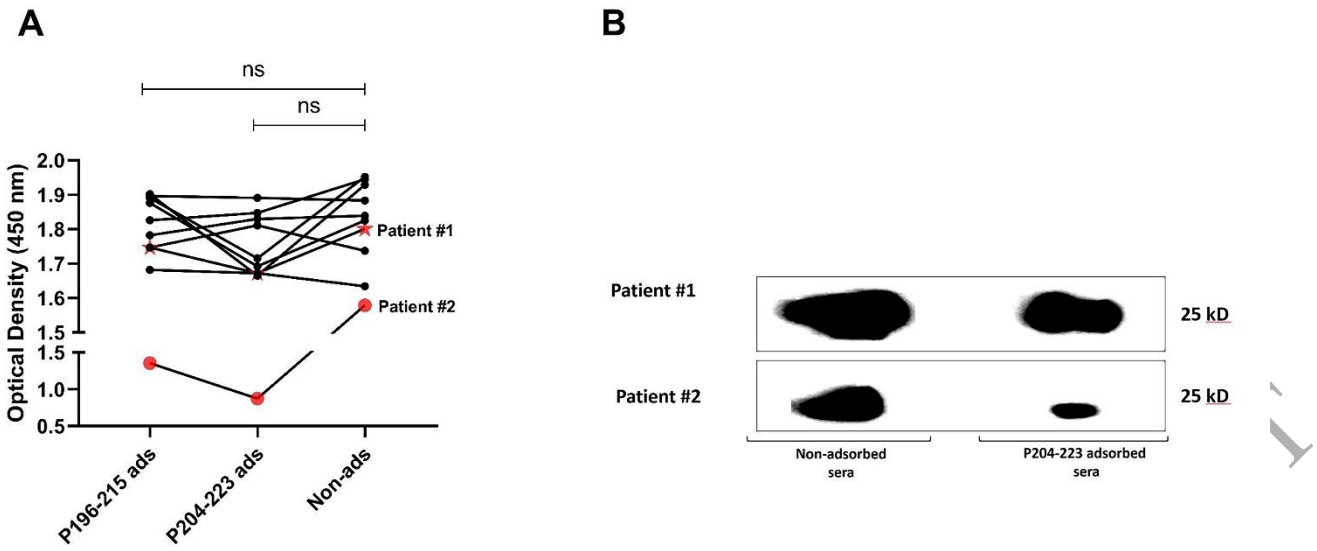

Fig 4. Depletion assay with P196-215 and P204-223 peptides of SARS-CoV-2 RBD protein. (A) Specific reactivity of 10 COVID-19 patients' sera antibodies to P196-215 and P204-223 peptides of RBD protein was assessed by evaluating the reactivity of P196-215 and P204-223 adsorbed and non-adsorbed sera against native RBD by ELISA. **(B)** Two representative serum samples (Patient #1 (star shape) and Patient #2 (bullet shape)) were also tested by Western blot on RBD protein (25 kD). RBD: Receptor binding domain.

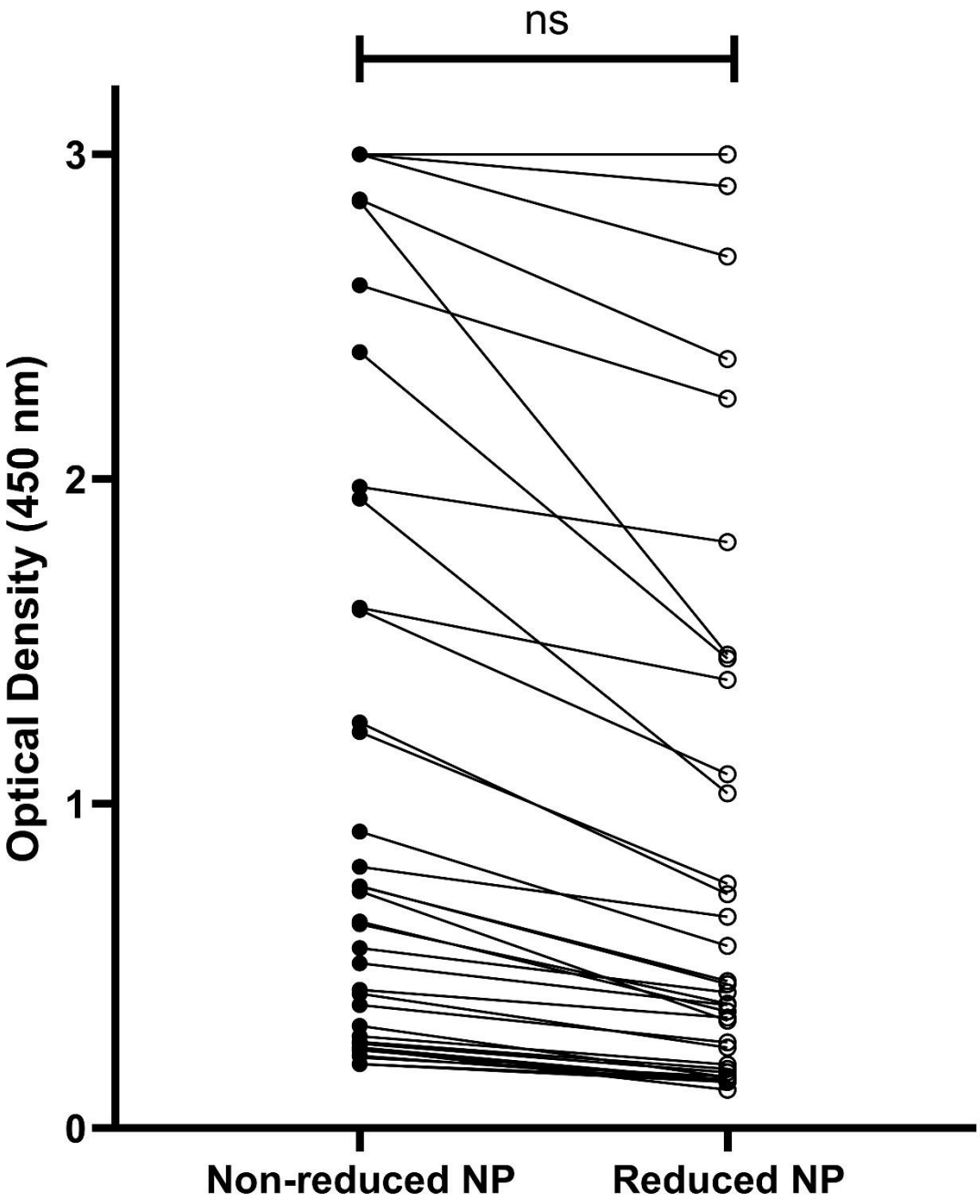

Fig 5. Reactivity of patients' sera with native and reduced SARS-CoV-2 NP protein. Native and reduced SARS-CoV-2 NP proteins were used for coating ELISA plates, and the sera from 48 COVID-19 patients were tested against each form of protein. The total reactivity of patients' sera with reduced NP was comparable with the corresponding reduced protein. NP: Nucleocapsid protein.

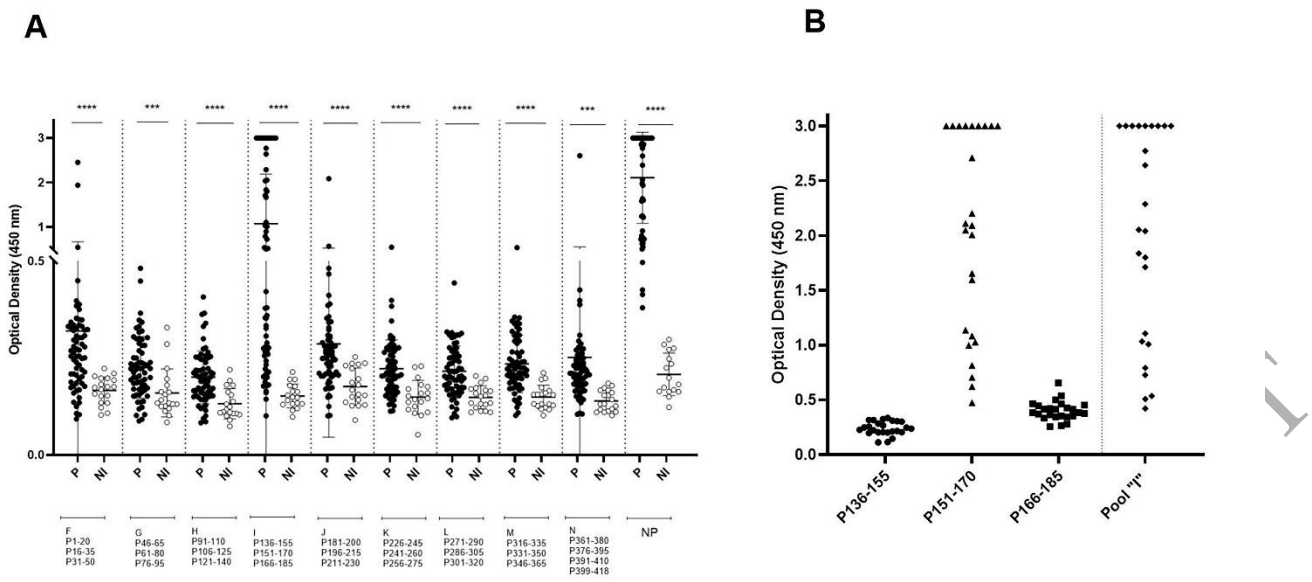

Fig 6. Mapping of the immunodominant epitopes on the SARS-CoV-2 NP protein by Pepscan analysis against antisera from COVID-19 patients and identification of the immunodominant epitope. (A) Sera of COVID-19 patients (n = 66) and healthy blood donors (n=23) were subjected to 9 pools (F to N) of overlapping peptides that cover the entire NP protein sequence. The reactivity of sera with the overlapping peptide pool that covers the entire sequences of the NP protein was determined by ELISA. **(B)** Identification of the immunodominant epitope among "pool I" peptides from NP protein. Individual peptides (P136-155, P151-170 and P166-185) of the positive peptide pool on NP protein was screened with plasma samples from 25 COVID-19 patients to determine the most reactive peptide. The solid lines represent mean values ± SD, *P<0.001*(***), *P<0.0001*(****). NP: Nucleocapsid; P: Patients; NI: Normal individuals.

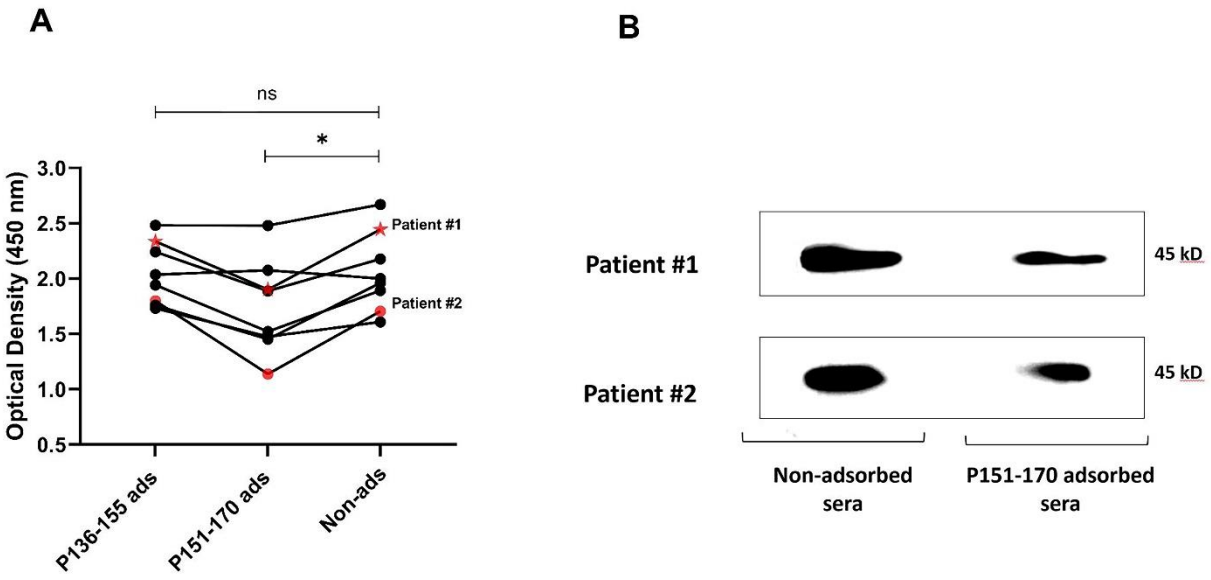

Fig 7. Depletion assay with P136-155 and P151-170 peptides of SARS-CoV-2 NP protein. (A) Reactivity of NP-specific antibodies from 10 patients to P136-155 and P151-170 peptides was assessed in P136-155 and P151-170 adsorbed and non-adsorbed sera against reduced and non-reduced forms of NP by ELISA. **(B)** Two representative serum samples (Patient #1 (star shape) and Patient #2 (bullet shape)) were also tested by Western blot on NP protein (45 kD). *P<0.05*(*). NP: Nucleocapsid.