Pak. J. Biochem. Mol. Biol., 2022, 55(2):138-145

Research paper

Recombinant production and validation of immunogenic fragment from RBD of SARS-CoV-2 spike protein

Sania Javaid, Zara Ashi, Aqsa Anwar, Mohsina Akhter, Shaista Arif, Mohsin Shad, M. Waheed Akhtar and Muhammad Sajjad*

School of Biological Sciences, University of the Punjab, Lahore-54590, Pakistan.

*Corresponding author: sajjad.sbs@pu.edu.pk

ABSTRACT

The rapid and widespread infection caused by SARS-CoV-2, a new type of coronavirus shook the whole world playing havoc with human life. Although the situation has been brought under control, the susceptibility level for its similar infection from coronavirus remains a high possibility. To safeguard from such situations the development of a reliable, economical, and high throughput method to diagnose the disease is required for preparedness in such situations. Serodiagnosis-detecting antibodies produced against antigens of SARS-CoV-2 in blood samples could prove beneficial. Based on the epitope position in the molecule, an immunogenic fragment from RBD of the S1 domain of spike protein of SARS-CoV-2; S1F2, was selected and expressed in Escherichia coli. Screening of S1F2 against 319 sera samples from COVID-19 patients showed a sensitivity of 41.07% and a corresponding specificity of 96.7%. Data regarding 3-D model prediction based on computational modeling was also in agreement. Solvent accessibility analysis through CPORT showed favorable orientation of the epitopic residues resulting in improved antigen-antibody interactions. This construct can be used further for producing constructs with greater sensitivity through the fusion of epitopes from additional SARS-CoV-2 antigens.

KEYWORDS Validation, Immunogenic, RBD, Spike protein, SARS-CoV-2

INTRODUCTION

Severe Acute Respiratory Syndrome (SARS-CoV-2); a new type of coronavirus, is known to cause the recent COVID-19 epidemic which occurred in December 2019. In a short time, the infectious disease spread throughout the world [1].WHO announced the COVID-19 a pandemic on 11th March 2020 [2]. So far above 546 million people have been infected with the virus and more than 6.3 million casualties have been reported globally as of July 17, 2022 [3]. Clinical studies have shown that mortality rates are higher in the elderly population [4] and much lower in children [5, 6].

The virus primarily infects the RT along with other organ systems [7]. It enters the cells by interacting with the cell surface receptor ACE2. It affects lung tissue causing severe damage to alveoli, which further can result in Acute Respiratory Distress Syndrome (ARDS). Along with the respiratory tract, the novel coronavirus also affects other organs of the human body including the GI tract, heart, kidney, liver, and brain [8].

Upon analyzing the clinical manifestations of SARS-CoV-2, it was discovered that the symptoms of the infection vary greatly among different individuals. Some patients present mild to severe symptoms e.g. respiratory discomfort, fever, fatigue, sore throat, and flu while some other patients

have no symptoms at all, known as asymptomatic [9].

The virus emerged in the first, second, and third waves over different time frames since the first case was detected in December 2019 [10]. Since no definitive therapy was available, the strategies of social distancing and lockdown were adopted to limit the spread of infection [11, 12]. The pandemic is not necessarily over yet; with over 5.7 million new cases reported globally in the first week of July [13]. As the virus spread, its variants were also reported including alpha, beta, gamma, and delta [14]. There is currently no targeted therapy available. Several drugs have been tested in clinical trials [15, 16], but none offers a definitive therapy [7].

A sensitive, cost effective and high throughput diagnostic test is essential for diagnosis and thus effective control of the disease. Serodiagnosis of high sensitivity can be instrumental in detection and control of COVID-19 pandemic. Serodiagnosis based on virus-specific antibodies is an effective approach for COVID-19 detection [17]. A variety of tests based on the principle of detecting antibodies produced against SARS-CoV-2 structural proteins (nucleocapsid and spike protein) have been developed commercially. Although such tests can be developed rapidly, in each case there is a mixture of virus-specific and non-specific epitopes which differ in antibody reactivity thus leading to compromised sensitivity and specificity [18].

At present time, several serodiagnostic methods are available for the detection of COVID-19 infection but all these tests retain low sensitivity and specificity. In this study, an immunogenic fragment from Receptor Binding Domain (RBD) of spike protein of SARS-CoV-2, which is unique to the virus, was produced by recombinant DNA technology and evaluated for its sensitivity for the detection of antibodies produced against COVID-19 in the blood samples.

MATERIALS AND METHODS

Bacterial strain and growth conditions

Two strains of *Escherichia coli* i.e. DH5α for cloning and BL21-CodonPlus (DE3)-RIPL for expression analysis, were used. Luria Bertani medium (LB) was used for bacterial growth.

Selection of the antigen

F2 fragment from Receptor Binding Domain (RBD) of spike protein of SARS-CoV-2 was used in this study (Fig. 1). The complete nucleotide sequence of spike protein of SARS-CoV-2 was obtained from GenBank(https://www.ncbi.nlm.nih.gov/genbank/). The construct was synthesized into the pUC57 cloning vector from SynBio Technologies.

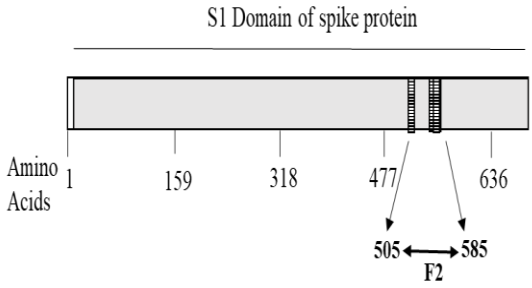

Figure 1: Schematic representation showing position of the fragment F2 in the S1 domain of spike protein

Cloning, Expression and Purification

F2 gene sequence was amplified by using S1-pUC57 (S1 domain of SARS-CoV-2 ligated with cloning vector pUC57). NEBcutter [19] , Primer3 [20], and OligoCalc [21] were used to design gene specific primers (Table 1).

The PCR conditions used for the molecule was initial denaturation at 95 °C for 5 min followed by 35 cycles, each consisting of denaturation at 95 °C for 45 s, annealing at 63 °C for 45 s, and extension at 72 °C for 45 s. The final extension was done at 72 °C for 10 min.

Table 1: Primers used for PCR amplification of F2 fragment.

Protein	Primer Name	Sequence (5′→ 3′)	R.E.	Tm (ºC)
F2 (spike)	S2F	5′-TATACATATGTACCAACCTTACAGAGTAG-3′	*Nde*I	64.8
	S2Rx	5′-TTAGAGCTCTCAAAGAATCTCAAGTGTCTG-3′	*Sac*I	68

The amplified gene product was ligated into pJET1.2/blunt cloning vector and conserved in *E. coli* DH5α by transformation with the recombinant DNA. The gene was then sub-cloned in pET28a(+) expression vector. The confirmation was done through colony PCR and restriction analysis. Combination of *Nde*I and *Sac*I was used for the restriction confirmation of F2 in recombinant plasmid.

After subcloning in pET28a(+), the antigen was expressed in *E. coli* BL21-CodonPlus (DE3)-RIPL. The transformed cells were grown in LB medium at 37 ºC until OD_{600} reached 0.6-0.8. Induction was carried out with 0.5 mM IPTG as an inducer and post induction incubation of 4 hours was carried out at 37 ºC. 12% SDS PAGE was used for expression analysis of F2. The cells expressing the protein of interest were harvested and resuspended in the buffer containing 50 mM Tris-Cl (pH 8.0), 0.2 M NaCl and 1 mM PMSF. The cells were then lysed by ultrasonication (UP400S ultraschallorozessor Dr. Hielscher GmbH, Teltow, Germany). The lysate was centrifuged at 6000 rpm for 10 minutes to separate the soluble and insoluble fractions. After sonication, total cell protein, soluble and insoluble fractions were analyzed on SDS PAGE. The percentage expression of protein was determined by densitometric analysis using SynGene gel documentation system (UK). The protein was expressed in insoluble form and was solubilized in 8M urea containing 50 mM Tris-Cl (pH 8.0), 0.2 M NaCl, and 1mM PMSF. The purification was done through Ni^{+2} affinity chromatography. The purified fractions containing protein of interest, as analyzed by SDS PAGE were pooled together and dialyzed against 50 mM Tris-Cl (pH 8.0).

Plasma sample collection

Plasma samples from 319 COVID-19 patients were collected from one of the national covid center at Jinnah Hospital Lahore. Human samples were collected from July 2020 to November 2021when the disease was at its peak. 163 plasma samples from the healthy individuals were also collected as the controls. Before initiating the experimental work, due approvals of SOPs for patient handling, sample collection and processing as well as consent of the patient from the concerned Ethical Review Board was also obtained. The informed consent was obtained in writing by filling the relevant performa.

Validation of the antigen

The immunoreactivity of the antigen for antibodies produced against SARS-CoV-2 antigens was checked through Enzyme Linked Immunosorbent Assay (ELISA). Polystyrene 96 well microtiter plates (Nunc, Denmark) were coated with 50 µL/well of antigen at a concentration of 3 µg/mL suspended in 1X PBS buffer (pH 7.4). Plates were incubated overnight in humid conditions at 4 ºC. Each microtiter plate was thoroughly washed thrice with 1X PBS supplemented with 0.05% tween 20 (1X PBS-T). 5% skim milk in 1X PBS (300 µL/well) was added as blocking agent, incubated at 37ºC for 2 hours and plates were washed as mentioned earlier. 50 µL of human plasma sample in 1:250 dilution (1X PBS + 5% skim milk) was then added to antigen coated plates and incubated at 37oC for 1 hour. After four washings with PBS-T, microtiter plates were further coated with 50 µL of anti-human IgG, conjugated with horseradish peroxidase at 1:10,000

dilution and incubated for 2 hours at 37°C. The plates were washed 3-4 times and reacted with TMB solution (50µL/well) for 5-10 minutes. Finally, 50 µL of 2M H2SO4 was added to stop the reaction. The blue color turned into yellow color after addition of stop solution. Differential absorbance for each plasma sample was recorded at 450nm/630nm using ELISA reader (HUMAREADER plus, human GMBH). All the sera samples were tested twice for their immunoreactivity with the recombinant protein and their average values were recorded.

Statistical analysis

ELISA results were analyzed by defining cutoff value, which helped in calculating sensitivity and specificity of the antigens. Cutoff value was calculated by the formula: Mean OD450/630 of HC plus 2.756 x SD of healthy control. All those samples were considered positive which had cutoff value above 1.0. All the O.D. values were normalized by dividing these with the relative cutoff value. Sensitivity was calculated by dividing number of positive samples by total number of patients being studied and specificity was calculated by dividing number of healthy individuals, having normalized OD 450/630 by the total number of healthy controls. Receiver operating characteristic (ROC) analysis was carried out using all the O.D. values to obtain area under the curve. Differences in healthy and patients were compared on non-parametric Mann-Whitney algorithm by considering results significantly reliable with $P<0.0001$. All the data was statistically analyzed in Graphpad Prism 5 software (Graph-Pad Software Inc., San Diego, CA).

Molecular modelling

3-dimensional structure of chimeric protein was predicted using Robetta server [22]. Moreover, solvent accessibility analysis of the predicted model was performed by CPORT [23] online server and structure was visualized using PyMOL (https://pymol.org/).

RESULTS

B-cell epitopes

The antigenic protein F2 was selected due to its high immunogenic potential. The epitope for F2 were predicted using the Bepipred-1.0 Linear Epitope Prediction software, which showed the presence of three B-cell epitopes.

Protein expression and purification

SDS PAGE analysis of *E. coli* cell proteins expressing the protein of interest showed the expected position of bands on gel with respect to size (Fig. 2-A). The expression level of S1F2 as percentage of the total cell protein was 30%. The protein was produced as inclusion bodies and was refolded into soluble form as described above. Purification done with Ni^{+2} chelating chromatography yielded protein with a purity of more than 90% (Fig. 2-B).

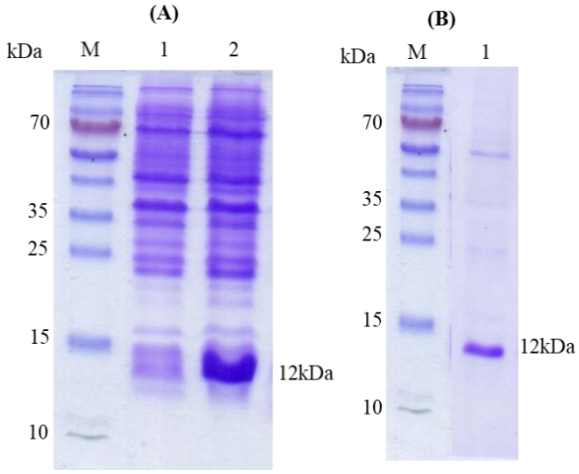

Figure 2: (A) SDS PAGE showing expression of the peptide S1F2. lane M: size markers; lane 1: proteins of the uninduced *E.coli* cells; lane 2: proteins of E. coli cells expressing S1F2. (B) SDS PAGE showing purified S1F2 protein. lane M: size markers; lane 1: partially purified S1F2

Sensitivity validation

On the basis of screening 319 plasma samples of COVID-19 patients, S1F2 detected 131 samples as positive with a sensitivity of 41.07% (Fig. 3-A). The specificity of S1F2 was 96.7%. Detection potential of the protein, calculated using Graphpad Prism 5 software, was recorded

as Receiver Operator Characteristic (ROC) curve as shown in Fig. 3-B. Value of Area Under Curve for S1F2 was 0.859.

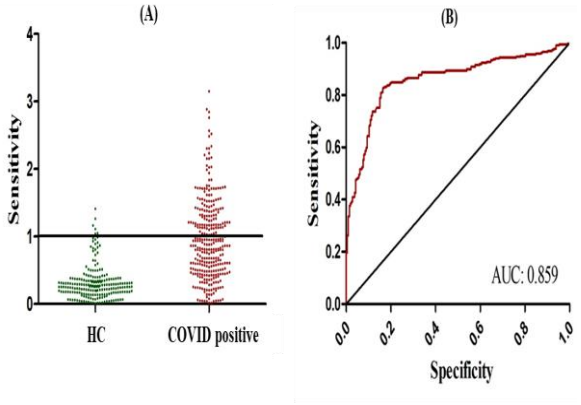

Figure 3: **Antibody detection through ELISA (A)** Scatter dot plot of normalized $OD_{450/630}$ showing plasma reactivity of 319 COVID-19 patients and 163 healthy controls with single antigen S1F2 **(B)** ROC curve showing diagnostic accuracy of ELISA for S1F2

Molecular modelling

3D molecular model of the protein was constructed using Robetta as shown in Fig. 4-A. Solvent accessibility of the constructed models was determined using CPORT server and visualized through PyMOL Fig. 4-B. Red color represented actively interacting residues along with supporting residues shown in green. Blue color represented residues that do not take part in any kind of interaction.

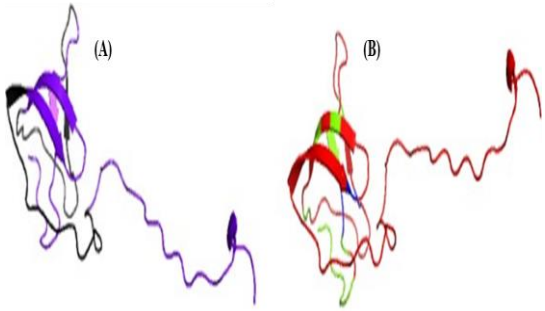

Figure 4: 3D molecular modelling (A) and analysis of solvent accessibility (B) of F2 gene fragment from spike protein of SARS-CoV-2. Epitopes in (A) are shown in black color whereas red color in (B) represents active residues with green color marked supporting residues for antibodies. Blue color in (B) represents the region that is not immunogenic.

DISCUSSION

The SARS-CoV-2 pandemic has emphasized the significance of laboratory diagnosis in health emergency management that has greatly influenced the global social, economic, and health fabric. There is a need to develop rapid and accurate serological test which will aid in pandemic management while being cost effective, time saving, and reducing burden in diagnostic laboratories and health-care settings.

The above mentioned objectives can be achieved by developing a molecule based on B-cell specific epitopes of SARS-CoV-2 antigen. The experimental work was planned on the basis of the structural protein S, which play a most dominant part in the structural and functional properties of the virus with respect to infection and immunological properties. S1 domain of the Spike protein has six reported epitopes [24]. As Spike protein has larger size so truncations were designed to expose the epitopes.

In this research work an immunogenic fragment from RBD of S1 domain was designed, which carry highly expressing immunogenic epitopes. Previously, in other sero-diagnostic tests full domains of Spike protein, RBD and their fusions was used [25, 26] but the use of epitopic region containing fragments unique to the virus favors the more appropriate detection of antibodies. As epitopes will be more exposed in the recombinant protein, so antibodies will be detected more efficiently even if present in minor quantity.

The fragment S1F2 was cloned in pET28a(+) and expressed in *E. coli* BL21-

CodonPlus (DE3)-RIPL strain. S1F2 showed percentage expression of about 30% [27]. The expression was optimized at 0.5 mM IPTG. The protein was expressed as inclusion bodies and was solubilized using 8M urea and was purified > 90% using Ni^{+2} affinity chromatography. Serodiagnostic potential of S1F2 was evaluated with the help of ELISA performed on the plasma samples from COVID-19 patients. The antigen when screened against 319 plasma samples of COVID-19 patients showed 41.07% sensitivity and a specificity of 96.7%.

Detailed *in silico* analysis of the fragment was done using the currently available softwares. *In silico* analyses were focused to the analysis of folding patterns and CPORT analysis for solvent accessibility. The data thus obtained showed that epitopes are in suitable position; represented in black color in 3D modelling and in red color through CPORT analysis, which are available for interacting with the antibodies and thus led to increased sensitivity.

This fragment S1F2, therefore, can be a potential component for developing a successful serodiagnostic procedure for SARS-CoV-2. Further, aim is to produce different recombinant fusion constructs of Spike Protein and Nucleocapsid protein with the best orientation of epitope exposure that can detect the antibodies IgG and IgM with greater efficiency by ELISA method. Main purpose of this experiment is to produce high throughput, more sensitive, more specific and cost effective sero-diagnostic kit.

Acknowledgment

The study is part of a Higher Education Commission (HEC) funded project under Rapid Research Program (RRG) – 2020 (No. RRG-43).

REFERENCES

1. Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W. and Lu, R. (2020). A novel coronavirus from patients with pneumonia in China, 2019. *New England journal of medicine*.

2. ur Rehman, M. F., Fariha, C., Anwar, A., Shahzad, N., Ahmad, M., Mukhtar, S. and Haque, M. F. U. (2021). Novel coronavirus disease (COVID-19) pandemic: A recent mini review. *Computational and Structural Biotechnology Journal*, **19**:612-623.

3. World Health Organization (2022). COVID-19 weekly epidemiological update, edition 99, 6 July 2022. last accessed on

4. Zhou, F., Yu, T., Du, R., Fan, G., Liu, Y., Liu, Z., Xiang, J., Wang, Y., Song, B. and Gu, X. (2020). Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *The lancet*, **395**:1054-1062.

5. Lu, X., Zhang, L., Du, H., Zhang, J., Li, Y. Y., Qu, J., Zhang, W., Wang, Y., Bao, S. and Li, Y. (2020). SARS-CoV-2 infection in children. *New England Journal of Medicine*, **382**:1663-1665.

6. Qiu, H., Wu, J., Hong, L., Luo, Y., Song, Q. and Chen, D. (2020). Clinical and epidemiological features of 36 children with coronavirus disease 2019 (COVID-19) in Zhejiang, China: an observational cohort study. *The Lancet infectious diseases*, **20**:689-696.

7. Yuki, K., Fujiogi, M. and Koutsogiannaki, S. (2020). COVID-19 pathophysiology: A review. *Clinical immunology*, **215**:108427.

8. Shah, M. D., Sumeh, A. S., Sheraz, M., Kavitha, M. S., Maran, B. A. V. and Rodrigues, K. F. (2021). A mini-review on the impact of COVID 19 on vital organs. *Biomedicine & pharmacotherapy*, **143**:112158.

9. Pan, X., Chen, D., Xia, Y., Wu, X., Li, T., Ou, X., Zhou, L. and Liu, J. (2020). Asymptomatic cases in a family cluster with SARS-CoV-2 infection. *The Lancet Infectious Diseases*, **20**:410-411.

10. Fisayo, T. and Tsukagoshi, S. (2021). Three waves of the COVID-19 pandemic. *Postgraduate medical journal*, **97**:332-332.

11. Farsalinos, K., Poulas, K., Kouretas, D., Vantarakis, A., Leotsinidis, M., Kouvelas, D., Docea, A. O., Kostoff, R., Gerotziafas, G. T. and Antoniou, M. N. (2021). Improved strategies to counter the COVID-19 pandemic: Lockdowns vs. primary and community healthcare. *Toxicology Reports*, **8**:1-9.

12. O'Dowd, A. (2020). Covid-19: Avoid indoor mixing over Christmas or risk third wave, warns iSAGE. British Medical Journal Publishing Group.

13. World Health Organization (2022). Weekly Epidemiological Update on COVID-19 - 13 July 2022. https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19---13-july-2022, last accessed on

14. Lauring, A. S. and Malani, P. N. (2021). Variants of SARS-CoV-2. *JAMA*, **326**:880-880.

15. Cao, B., Wang, Y., Wen, D., Liu, W., Wang, J., Fan, G., Ruan, L., Song, B., Cai, Y. and Wei, M. (2020). A trial of lopinavir–ritonavir in adults hospitalized with severe Covid-19. *New England Journal of Medicine*.

16. Gautret, P., Lagier, J.-C., Parola, P., Meddeb, L., Mailhe, M., Doudier, B., Courjon, J., Giordanengo, V., Vieira, V. E. and Dupont, H. T. (2020). Hydroxychloroquine and azithromycin as a treatment of COVID-19: results of an open-label non-randomized clinical trial. *International journal of antimicrobial agents*, **56**:105949.

17. Burbelo, P. D., Riedo, F. X., Morishima, C., Rawlings, S., Smith, D., Das, S., Strich, J. R., Chertow, D. S., Davey Jr, R. T. and Cohen, J. I. (2020). Detection of nucleocapsid antibody to SARS-CoV-2 is more sensitive than antibody to spike protein in COVID-19 patients. *MedRxiv*.

18. Gomes, L. R., Durans, A. M., Napoleao-Pego, P., Waterman, J. A., Freitas, M. S., De Sa, N. B., Pereira, L. V., Furtado, J. S., Aquino, R. G. and Machado, M. C. (2021). Multiepitope proteins for the differential detection of IgG antibodies against RBD of the spike protein and non-RBD regions of SARS-CoV-2. *Vaccines*, **9**:986.

19. Vincze, T., Posfai, J. and Roberts, R. J. (2003). NEBcutter: a program to cleave DNA with restriction enzymes. *Nucleic acids research*, **31**:3688-3691.

20. Kõressaar, T., Lepamets, M., Kaplinski, L., Raime, K., Andreson, R. and Remm, M. (2018). Primer3_masker: integrating masking of template sequence with primer design software. *Bioinformatics*, **34**:1937-1938.

21. Kibbe, W. A. (2007). OligoCalc: an online oligonucleotide properties calculator. *Nucleic acids research*, **35**:W43-W46.

22. Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G. R., Wang, J., Cong, Q., Kinch, L. N. and Schaeffer, R. D. (2021). Accurate prediction of protein structures and interactions using a three-track neural network. *Science*, **373**:871-876.

23. de Vries, S. J. and Bonvin, A. M. (2011). CPORT: a consensus interface predictor and its performance in prediction-driven docking with HADDOCK. *PloS one*, **6**:e17695.

24. Yang, X., Yu, Y., Xu, J., Shu, H., Liu, H., Wu, Y., Zhang, L., Yu, Z., Fang, M. and Yu, T. (2020). Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. *The Lancet Respiratory Medicine*, **8**:475-481.

25. Thomas, A., Messer, W. B., Hansel, D. E., Streblow, D. N., Kazmierczak, S. C., Lyski, Z. L., Lu, Z. and Slifka, M. K. *Establishment of monoclonal antibody standards for quantitative serological diagnosis of SARS-CoV-2 in low-incidence settings*. in *Open Forum Infectious Diseases*. 2021. Oxford University Press US.

26. Perera, R. A., Mok, C. K., Tsang, O. T., Lv, H., Ko, R. L., Wu, N. C., Yuan, M., Leung, W. S., Chan, J. M. and Chik, T. S. (2020). Serological assays for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), March 2020. *Eurosurveillance*, **25**:2000421.

27. Arif, S., Akhter, M., Khaliq, A., un Nisa, Z., Khan, I. H. and Akhtar, M. W. (2021). Serodiagnostic evaluation of fusion proteins from multiple antigens of *Mycobacterium tuberculosis* for active TB. *Tuberculosis*, **127**:102053.

Pak. J. Biochem. Mol. Biol., 2022, 55(2):138-145