Pak. J. Biochem. Mol. Biol., 2022, 55(3):148-155

Research paper

Potential of an Epitope-rich Peptide from S2 Domain of Spike Protein in Serodiagnosis of SARS-CoV-2

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

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

*Corresponding author: Prof. Dr. M. Waheed Akhtar (mwa.sbs@pu.edu.pk)

ABSTRACT

A rapid and highly reliable diagnostic method is paramount to control spread of the life threatening COVID-19. Immunogenic fragment rich in epitopes could provide high sensitivity and specificity for detecting antibodies in serodiagnosis of SARS-CoV-2. Based on epitope positions on spike protein a fragment F5 having fourteen known epitopes, was generated from S2 domain of spike protein. F5 was cloned in pET28a (+) using restriction sites NheI and XhoI. The recombinant plasmid was expressed in E. coli BL21-CodonPlus (DE3)-RIPL inducing the cells with the optimized concentration of 0.1 mM IPTG. Protein inclusion bodies were solubilized in 8 M Urea. Solubilized protein was purified by Ni+2 affinity chromatography and urea content was removed by dialysis. The antigenic fragment F5 thus produced, could be used to raise antisera for applications in diagnostic procedures based on ELISA. Being rich in epitopes, serodiagnosis based on this immunogenic fragment can provide high sensitivity and specificity of neutralization effect against anti- SARS-CoV-2 antibodies in COVID-19 patients.

KEYWORDS SARS-CoV-2, COVID-19, epitope, S2 domain, spike protein

INTRODUCTION

During the past two decades, the world has experienced outbreaks of respiratory syndromes caused by coronaviruses [1,2]. The more recent syndrome COVID-19 occurred as a pandemic with a high fatality. Early cases of COVID-19 were reported in Huanan market in Chinese city Wuhan with Pneumonia like symptoms caused by a novel coronavirus, which was named as 2019-nCoV by Center of Disease Control in China. It spread quickly causing severe symptoms and the World Health Organization declared it as pandemic and renamed 2019-nCoV as SARS-CoV-2 (severe acute respiratory syndrome coronavirus-2) [3].

SARS-CoV-2 is a pathogenic, single stranded, positive sense and polyadenylated RNA virus [4]. It spreads mainly from breathing in the droplets or aerosols in the air that come from the mouth and nose of infected person. The symptoms of COVID-19 like fever, fatigue, cough or myalgia are very similar to those of the other viral infections like influenza, making its lab based diagnosis necessary [5]. Currently used diagnostic methods include RT-PCR (Real time polymerase chain reaction), CT imaging (computerized tomography imaging) and those based on hematological parameters. RT-PCR has become the standard diagnostic tool for COVID-19 diagnosis. But it has some limitations as it takes 2-3 hours to generate results, require highly trained technicians, expensive equipment and also give a number of false negative results [6]. Sensitivity of the method completely rely on concentration

of RNA causing limitations for a simple and rapid diagnosis for screening COVID-19 patients [7].

Assays based on antibody detection for the detection of SARS-CoV-2 have advantages over the PCR-based test. Sample collection of sera is easier and safer than the swab specimen collection for RT-PCR and it also skips the highly sensitive RNA isolation steps [8]. Different techniques are being used for serodiagnostic purpose like Point of care (POC) lateral flow immunoassays (LIFAs), Chemiluminescent assay and enzyme linked immunosorbent assay (ELISA) [9]. The more common of these is the enzyme linked immunosorbent assay (ELISA), where a plate coated with viral antigens is exposed to the patient blood carrying the antibodies. The complex of antigen and antibody is detected by a labelled secondary antibody and a visual color appearance indicates the presence of antibodies (IgG, IgA, and IgM) [10]. Currently available serodiagnostic methods mostly uses full length spike protein or nucleocapsid as antigen and have low level of sensitivity and specificity [11].

This study focuses on developing a reliable serodiagnostic method in comparison to the recently used diagnostic procedures by using an epitope rich fragment from spike protein, which is highly immunogenic and could provide high sensitivity and specificity against anti-SARS-CoV-2 antibodies.

MATERIALS AND METHODS

Selection of the fragment from spike protein

Potential epitope regions of spike protein were retrieved by using immune epitope database (IEDB). Out of these an epitope rich fragment F5 in S2 domain was selected for this study. Further potential epitopes in this fragment were predicted using Bepipred-1.0 (http://tools. iedb.org/bcell), which is an online tool used for prediction of linear B-cell epitopes based on hidden Markov rule and propensity scale method [12]. Further these epitopic regions were confirmed by using Bepipred-2.0 which uses random forest algorithm for sequential B-cell epitope prediction [13].

PCR amplification

Nucleotide sequence of S2 domain was retrieved from NCBI GenBank and construct S2-pUC57 was commercially synthesized from Macrogen, South Korea. S2-pUC57 was used as template to amplify the F5 gene fragment by using the primers S5F and S5Rx. Primers of the gene fragment F5 were designed using online tool; Primer 3. Forward primer of F5 (SF5) had restriction site of *Nhe*I (5′ GCTAGC 3′) and reverse primer (S5Rx) had restriction site of *Xho*I (5′ CTCGAG 3′) [Table 1].

Table 1: Primers for PCR amplification of F5

Primer Name	Sequence 5'-3'	GC Content (%)	Restriction site	Tm (°C)
S5F	TTTAGCTAGCGTTTTGCCACCTTTGC	46	*Nhe*I	66.2
S5Rx	GTAACTCGAGTTATACACCATGAGGTGCTG	47	*Xho*I	61.05

Cloning of F5 in pJET1.2/blunt end cloning vector

PCR amplicons were purified from gel using GeneJET™ Gel Extraction Kit and blunt end ligation was used to clone the PCR product into pJET1.2/blunt end cloning vector. *E. coli* DH5α cells were transformed with plasmid pJET1.2-F5. pJET1.2/blunt end cloning vector has a unique screening method of positive clones as the insert is being ligated in a lethal gene disrupting the formation of this lethal protein that otherwise will not allow the growth of non-recombinant clones. Further colony PCR was used to confirm the presence of recombinant plasmid in *E. coli* DH5α cells. Recombinant plasmid pJET1.2-F5 was isolated from positive

clones of *E. coli* DH5α using alkaline lysis method and the presence of F5 gene fragment in vector was confirmed by double digestion using *Nhe*I and *Xho*I as restriction enzymes. The restricted F5 gene fragment was ligated in expression vector pET28a(+) at the restriction sites of *Nhe*I and *Xho*I. Competent cells of *E. coli* DH5α were transformed with recombinant vector pET28a-F5 and the positive clones were confirmed both by colony PCR and restriction analysis of the recombinant plasmid pET28a-F5 (using restriction enzymes *Nhe*I and *Xho*I).

Protein expression and its optimization

F5 fragment was expressed in *E. coli* BL21-CodonPlus (DE3)-RIPL. LB medium was inoculated with *E. coli* BL21 colony having recombinant plasmid pET28a-F5 and incubated at 37 °C in an orbital shaker until OD_{600} reached 0.6-0.8. Aliquoted 1ml of uninduced culture and the remaining culture was induced with varying final concentration of IPTG i.e. 0.1 mM, 0.25 mM, 0.5 mM, 0.75mM and 1mM in different falcon tubes. Incubated the induced culture for 6 hours in orbital shaker at 37 °C. harvested the cells and analyzed both uninduced and induced samples by SDS-PAGE. For time optimization of protein expression culture was induced with 0.5 mM IPTG and 1ml of sample was aliquoted after every 2 hours. So, the samples with IPTG induction of 2 hrs, 4 hrs, 6 hrs, 8 hrs and 10 hrs was analyzed on SDS-PAGE. Expression of protein was also estimated by another inducer as lactose. Different concentrations of 0.5 M lactose were used for the induction i.e. 5 mM, 10 mM, 15 mM and 20 mM lactose at its final concentration.

Lysis of *E. coli* cells by Ultrasonication

Harvested cells of *E. coli* BL21 (containing recombinant plasmid pET28a-F5) were resuspended in 50mM Tris-Cl buffer (pH 8.0) and added 100 µl of 1 mM PMSF (dissolved in isopropanol). The cells were lysed by ultrasonication. Aliquoted the lysed sample for total cell protein analysis, and centrifuged the remaining to obtain the supernatant containing the soluble proteins and the insoluble pellet.

Solubilization and refolding the expressed peptide

As F5 fragment was expressed in the form of inclusion bodies, so urea was used as denaturant to solubilize the protein. Tris buffer with urea concentrations of 2 M, 4 M, 6 M and 8 M were used to optimize the solubility of inclusion bodies. As the recominant peptide F5 was expressed with the poly histidine affinity tag Ni^{+2} affinity chromatography was used for the purification of the peptide. Resin was equilibrated with 20 ml of equilibration buffer consisting of 20 mM Tris-Cl buffer pH (8.0), 0.3 N NaCl, 10 mM imidazole and 8 M urea. The peptide sample dissolved in 8 M urea was loaded on the nickel-resin cloumn and allowed to bind with resin for about 30 minutes. Unbound proteins were washed with 10 ml of washing buffer containing 20 mM Tris-Cl buffer, 0.3N NaCl, 20 mM imidazole and 8 M urea. 6His-tagged peptides were eluted with a step gradient of imidazole ranging in concentration of 50 mM - 500 mM in the Tris buffer . The fractions of 3 ml each thus obtained were analysed by SDS-PAGE. Bradford assays were done using bovine serum albumin as a standard (14)

Molecular modeling of F5 fragment

3-D homology model of recombinant F5 fragment was determined by Robetta server [15], which is template based modeling tool with improved alignment accuracy, also provides B-cell specific epitope regions in 3D model.

RESULTS

Cloning of the B-cell specific peptide F5

The peptide F5 selected for this study lies in the S2 domain of the spike protein of SARS-CoV-2. It consisted of the residues 860-1060 and is found to contain fourteen epitopes as predicted using the software Bepipred-1.0. F5 fragment (603 bp) was

amplified by polymerase chain reaction using S2-pUC57 (gene of S2 domain of Spike protein ligated with cloning vector pUC57) as template. 1% agarose gel was run for the analysis of amplified products and a sharp band at nearly 603 bp confirmed the amplification of F5 gene fragment [Fig. 1a]. PCR amplicons were gel purified, ligated into pJET1.2/blunt end cloning vector and sub-cloned into pET28a(+). These recombinant plasmids were transformed in *E.coli* DH5α and the presence of insert in the recombinant plasmid was confirmed by double digestion using restriction enzymes *Nhe*I and *Xho*I [Fig. 1b].

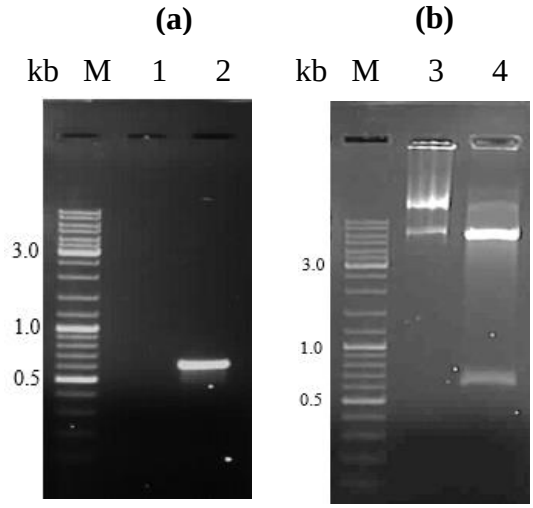

Figure 1: Agarose gel electrophoresis of PCR amplified F5 fragment (a) and the restriction digest of the recombinant pET28a-F5 (b). lanes M: size markers; lane 1: negative control for PCR and lane 2: amplified F5 gene (603 bp with overhangs); lane 3: uncut plasmid pET28a-F5; lane 24: double digest of the plasmid.

Optimization of Expression

After transforming the *E. coli* BL21 CodonPlus (DE3)-RIPL cells with the recombinant plasmid pET28a-F, the resultant colonies were cultivated in shake flasks and induced with 0.5 mM IPTG. SDS-PAGE analysis confirmed expression of the peptide of the expected size i.e., ~23 kDa. Type and concentration of inducer and time of maximum expression were also optimized. Expression of F5 fragment was examined by using the lactose as inducer at its varying concentrations i.e. 5 mM, 10 mM, 15 mM and 20 mM. SDS PAGE analysis revealed that final concentration of 10 mM lactose provided highest expression of the peptide. Expression of F5 fragment was also optimized by inducing the culture with 0.1 mM, 0.25 mM, 0.5 mM, 0.75 mM and 1 mM final concentration of IPTG. 0.1 mM IPTG was found to be sufficient for the best production of F5 fragment. Comparing the expression of lactose induction with expression obtained by IPTG, it was obvious that IPTG expressed high yield of protein even at its low concentration. So, IPTG was selected as inducer for the expression of F5 fragment.

Analysis of the expressed peptide was done at 2hrs, 4hrs, 6hrs, 8hrs and 10hrs of post induction for optimization of time for F5 expression. 4 hours of induction showed highest expression of the peptide. Thereafter the peptide expression level remained the same until after 8 hours the peptide level began to reduce. Although 4 hours of induction is sufficient for optimal protein production but 6 hours were selected for a higher number of cells.

Thus, the transformed *E. coli* BL21 were cultivated under the optimized conditions for expression of the peptide. The cells obtained by centrifugation were washed with Tris buffer. The harvested cells were then lysed by sonication and centrifuged to separate the soluble supernatant from the insoluble pellet. SDS-PAGE analysis of the fractions showed that peptide F5 appeared in insoluble cellular fraction. The formation of inclusion bodies of F5 peptide could be due to the presence of large number of hydrophobic amino acid residues.

Solubilization and purification of the expressed peptide

The inclusion bodies in the insoluble cell pellet were washed free from the cell debris by rinsing with the Tris buffer and

centrifugation at 500 x *g*. The pelleted inclusion bodies were dissolved in Tris buffer containing 8M urea [Fig 2a], and purified through Ni^{+2} affinity column. Analysis of the fractions obtained by SDS-PAGE showed that major fraction of the peptide was eluted with 150 mM imidazole concentration [Fig 2b]. Fractions with high concentration of F5 were pooled and dialyzed against 20mM Tris-Cl and a gradient concentration of urea. The purified F5 was analyzed on SDS-PAGE as shown in Fig. 2c. The sample thus obtained was quantified by bradford assay and it was found to contain 0.3 mg/ml.

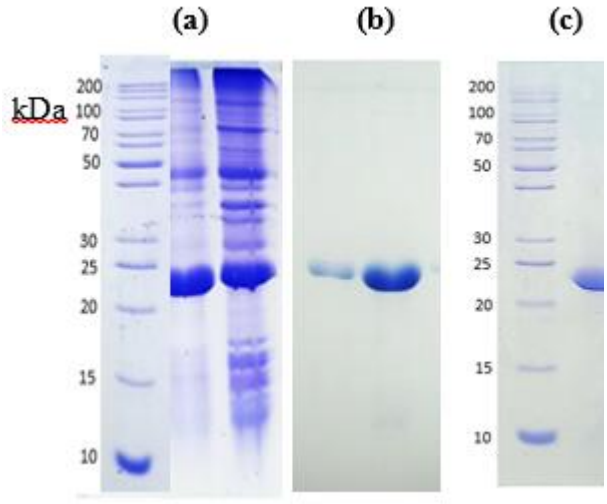

Figure 2: Expression of the peptide F5 in *E. coli* (a), purification of the solubilized F5 through nickel-affinity column (b) and the purified F5 (c). (a) lane M: size markers; lane 1: solubilized sample of F5: lane 2: total *E. coli* cell proteins expressing F5; lanes 3, 4: fractions eluted with imidazole from the nickel column; lane 6: purified F5.

In silico Structural Analysis

The primary structure and all the physiochemical properties of the antigenic fragment F5 werte analyzed by Expasy ProtParam tool. 3-dimensional model of F5 was constructed through Robetta server. This model shows alpha helices and beta sheets along with B-cell specific epitopes in immunogenic fragment F5 (Fig. 3). The figure shows the well exposed position of the epitopic regions in black colour in the peptide. Physical properties of the peptide as predicted by ExPASy ProtParam are summarised in Table 2.

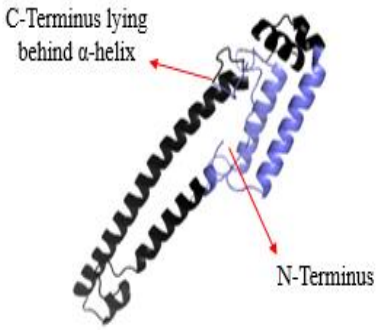

Figure 3: 3D model of F5 fragment constructed through Robetta server; Black color shows B-cell specific epitopes while blue color represents non-epitopic regions in F5 antigenic fragment. N and C termini are also labeled.

DISCUSSION

As pandemic of SARS-CoV-2 is causing havoc all over the world, there is an emergent need of early diagnostic and therapeutic techniques. Many molecular diagnostic methods such as nucleic acid sequencing, RNA amplification through RT-PCR or LAMP, Biosensor or CRISPR Cas-9 based techniques are currently in use [16]. While for serodiagnostic techniques, immunogenic proteins of SARS-CoV-2 such as Nucleocapsid and Spike protein are the potential candidates for diagnosis of COVID-19.N protein is less glycosylated that make it more immunogenic and it is tremendously conserved protein of coronaviruses. But serodiagnosis using N protein has been reported with false positive results as N protein has high morphological similarity with other HCoV proteins, so it may give neutralizing effects against other human coronaviruses like common cold viruses, HCoV-229E, HKU1 etc. [17].

Table 2: Physical properties of the F5 fragment predicted by ExPASy ProtParam.

Physical properties	**Values**
Amino acid residues	201
Molecular weight (kDa)	21.6
pI	8.60
Ext. coefficient (M^{-1} cm^{-1})	13075
Aliphatic index	101.99
Instability index	35.14
Grand average of hydropathicity (GRAVY)	0.099
Estimated half-life	>10 hours

Computational studies reveal that out of the 27 proteins of the SARS-CoV-2, 26 proteins have antigenic potential but spike protein S is the best candidate as it has a signal peptide, moderate aliphatic index, negative GRAVY value, long half-life, one transmembrane helix, beta wrap motifs as well as high stability [18]. Hence most of the serodiagnostic and therapeutic techniques use spike protein as target antigen. In spike protein, S1 domain is highly glycosylated while S2 domain is irregularly glycosylated within the whole region so it is less glycosylated and more immunogenic to antibodies as compared to S1 domain [19].

The focus of this study is that instead of using full length antigenic proteins as target for serodiagnostic or therapeutic purposes, only the rich epitopic regions could be used as immunogenic fragments. As full length viral proteins can also give ambiguous results due to cross-reactivity with phylogenetically similar viruses i.e. spike protein of SARS-CoV-2 has 90% sequence similarity with SARS-CoV spike protein [20]. To date, many other research works have been done based on serodiagnosis using full length spike protein, its domains or their fusions but the use of epitopic immunogenic fragments and their fusions would provide a high level of sensitivity and specificity [19].

Spike protein has S1 and S2 domains with many epitopic regions as described in section 2,1. Out of these epitopic regions, a fragment F5 ranging from 860-1060 amino acid residues residing in S2 domain was cloned and expressed. F5 has size of 603 bp gene fragment which in turn encodes protein with size of 21.6 kDa having structural and functional role in mediating virus-cell fusion and integrating it into host cell. F5 was expressed by using only 0.1 mM IPTG in E. coli BL21 cells, up to an expression level of 30%. Purification of F5 was performed through Ni+2 affinity chromatography. The dialyzed and purified protein was found to have a purity level more than 90%.

Structural analysis of the peptide F5 obtained using the different software tools show that the orientation of the epitopic residues as well as the physical parameters of the F5 fragment are favorable for using in combination with other peptides to develop a highly sensitive serodiagnostic probe for SARS CoV-2. F5 could be further used for antisera production and its detection through ELISA. It has fourteen epitopes, hence a high level of sensitivity for neutralizing antibodies could be expected from this immunogenic fragment. Using different epitope rich immunogenic fragment either in cocktail or by fusing them together in a single peptide would provide an elevated sensitivity level against anti-SARS-CoV-2 antibodies [21]. Other than serodiagnosis, these immunogenic fragments are also a suitable target for therapeutic purposes.

REFERENCES

1. Burrell, C.J., C.R. Howard, and F.A. Murphy, Coronaviruses. Fenner and white's medical virology, 2017: p. 437.

2. Abduljalil, J., Laboratory diagnosis of SARS-CoV-2: available approaches and limitations. New microbes and new infections, 2020. 36: p. 100713.

3. ur Rehman, M.F., et al., Novel coronavirus disease (COVID-19)

pandemic: A recent mini review. Computational and Structural Biotechnology Journal, 2021. 19: p. 612-623.

4. Astuti, I., Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2): An overview of viral structure and host response. Diabetes & Metabolic Syndrome: Clinical Research & Reviews, 2020. 14(4): p. 407-412.

5. Jackson, C.B., et al., Mechanisms of SARS-CoV-2 entry into cells. Nature Reviews Molecular Cell Biology, 2022. 23(1): p. 3-20.

6. Li, Z., et al., Development and clinical application of a rapid IgM-IgG combined antibody test for SARS-CoV-2 infection diagnosis. Journal of medical virology, 2020. 92(9): p. 1518-1524.

7. Corman, V.M., et al., Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Eurosurveillance, 2020. 25(3): p. 2000045.

8. Tao, K., P.L. Tzou, and J. Nouhin, The biological and clinical significance of emerging SARS-CoV-2 variants. 2021. 22(12): p. 757-773.

9. Poljak, M., et al., Clinical evaluation of the cobas SARS-CoV-2 test and a diagnostic platform switch during 48 hours in the midst of the COVID-19 pandemic. Journal of Clinical Microbiology, 2020. 58(6): p. e00599-20.

10. Alharbi, S.A., et al., Enzyme-Linked Immunosorbent Assay for the Detection of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) IgM/IgA and IgG Antibodies Among Healthcare Workers. Cureus, 2020. 12(9): p. e10285-e10285.

11. Long, Q.-x., et al., Antibody responses to SARS-CoV-2 in COVID-19 patients: the perspective application of serological tests in clinical practice. MedRxiv, 2020.

12. Larsen, J.E.P., O. Lund, and M. Nielsen, Improved method for predicting linear B-cell epitopes. Immunome research, 2006. 2(1): p. 1-7.

13. Jespersen, M.C., et al., BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. Nucleic Acids Res, 2017. 45(W1): p. W24-w29.

14. Bradford, M.M., A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Analytical biochemistry, 1976. 72(1-2): p. 248-254.

15. Kim, D.E., D. Chivian, and D. Baker, Protein structure prediction and analysis using the Robetta server. Nucleic Acids Res, 2004. 32(Web Server issue): p. W526-31.

16. Arumugam, A., et al., A rapid SARS-CoV-2 RT-PCR assay for low resource settings. Diagnostics, 2020. 10(10): p. 739.

17. Yamaoka, Y., et al., Whole Nucleocapsid Protein of Severe Acute Respiratory Syndrome Coronavirus 2 May Cause False-Positive Results in Serological Assays. Clin Infect Dis, 2021. 72(7): p. 1291-1292.

18. Can, H., et al., In silico discovery of antigenic proteins and epitopes of SARS-CoV-2 for the development of a vaccine or a diagnostic approach for COVID-19. Scientific Reports, 2020. 10(1): p. 22387.

19. Khare, S., et al., Conformational Changes of the Receptor Binding Domain of SARS-CoV-2 Spike Protein and Prediction of a B-Cell Antigenic Epitope

Using Structural Data. Frontiers in Artificial Intelligence, 2021. 4.

20. Molaei, S., et al., The immune response and immune evasion characteristics in SARS-CoV, MERS-CoV, and SARS-CoV-2: Vaccine design strategies. International immunopharmacology, 2021. 92: p. 107051.

21. Chan, A.P., Y. Choi, and N.J. Schork, Conserved Genomic Terminals of SARS-CoV-2 as Co-evolving Functional Elements and Potential Therapeutic Targets. bioRxiv : the preprint server for biology, 2020: p. 2020.07.06.190207.