Vaccine Research

ISSN: 2423-4923
eISSN: 2383-2819

Novel Protein Expression and Purification of SARS CoV-2 from Recombinant *Escherichia coli* System

Maryam Mashhadi Abolghasem Shirazi [1, 2], Mina Hannan [1], Golnaz Bahramali [1], Mohammad Reza Aghasadeghi [1], Seyed Mehdi Sadat [1]*

[1]Department of Hepatitis and AIDS and Blood borne diseases, Pasteur Institute of Iran, Tehran, Iran. [2] Molecular Virology Department, Pasteur Institute of Iran, Tehran, Iran.

ARTICLE INFO

Research Article

VacRes, 2023
Vol. 10, No.1, 6 – 10
Received: February 29, 2023
Accepted: March 19, 2023
Pasteur Institute of Iran

*** Corresponding Author:**
Seyed Mehdi Sadat;
69 Pasteur Ave., Hepatitis, AIDS and Blood-Borne Viruses Department, Pasteur Institute of Iran, Tehran1316943551, Iran.
Email: mehdi_sadat@pasteur.ac.ir
Tel /Fax: +98-21-66969291

KEYWORDS: SARS-CoV-2, Recombinant, Protein, *E. coli*, pET28a, Vaccine.

ABSTRACT

Introduction: The SARS-CoV-2 epidemic has infected and killed millions of people worldwide since 2019. Therapies available for SARS-CoV-2 disease are limited due to the sudden emergence of the virus. The need for appropriate vaccines to combat this dangerous virus is essential. Induction of specific antibodies against SARS-CoV-2 virus is mainly by nucleocapsid (N), membrane (M) and spike glycoprotein (S) proteins. On the other hand, cost-effective and fast methods for the expression and purification of recombinant proteins that retain antigenic properties are essential for vaccine development. **Methods:** A DNA fragment encoding the N, M and S1 proteins was inserted into pET28a vector to make an expression plasmid. The recombinant protein was then expressed in *Escherichia coli* Rosetta strain and purified by Ni-NTA column, followed by confirmation by SDS-PAGE and Western blotting. **Results:** PCR results showed that the gene was inserted correctly into the expression vector. The expression of the recombinant protein containing N, M and S1 proteins from SARS CoV-2 was optimized in a bacterial system. The recombinant protein was successfully purified from Ni-NTA column with a high yield **Conclusion:** Upon further evaluations, this cost-effective approach for the production of recombinant antigenic proteins in *E. coli* (Rosetta), could potentially be used for the development of vaccines against coronaviruses infections.

INTRODUCTION

Patients with coronavirus disease of 2019 (COVID-19) were referred to be infected with an acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [1]. According to the World Health Organization (WHO), the ensued pandemic caused more than 500 million confirmed cases and more than 6 million deaths by June 2022 [2, 3]. Vaccines have been researched and developed with different approaches to combat COVID-19 pandemic; therefore to date, more than 11 billion doses of vaccines have been used to immunize humans worldwide [3]. Unfortunately, the emergence of SARS-CoV-2 types has reduced the effectiveness of most vaccines due to multiple mutations in their surface proteins. Therefore, one of the mandatory measures to deal with different types of SARS-CoV-2 is to improve the immunogenicity of the vaccine [4]. A suitable approach to improve immunogenicity of a vaccine against COVID-19 could be addressed by optimizing its antigen design.

SARS-CoV-2 structural protein genes include surface spike glycoprotein (S), small membrane protein (M), envelope glycoprotein (E), and nucleocapsid protein (N) [5]. The N protein is involved in assembling the virus while the M and E proteins are involved in the formation of the virus coating [6]. N proteins have been shown to be potent in several immunogenic coronaviruses, such as murine coronavirus and respiratory syndrome coronavirus [5]. Most of the COVID-19 patients have shown a specific immune response against the full length and fragments of the N protein [7]. Among the proteins expressed by SARS-CoV-2, viral glycoprotein S is widely recognized as a therapeutic or vaccine target [8]. Coronaviruses enter the cell through the S-membrane glycoprotein which also protrudes from the surface of the virus. Glycoprotein S is made up of two subunits, called S1 at the amino terminus (responsible for binding to the target cell surface receptor) and S2 at the carboxyl terminus (responsible for the fusion of the virus and the cell membranes) [9, 10]. SARS-CoV-2 uses the angiotensin-converting enzyme 2 (ACE2) as a receptor to enter the cells. To enter a cell, the virus first binds to its ACE2 receptor in the RBD region, and following binding, spatial changes activate the S2 subunit which integrate with the cell membrane and let the viral particles to enter into the cell [11, 12]. The S1 subunit has two amino terminals (NTD) and a second receptor binding (RBD). It has a motif called the receptor-binding motif (RBM) within the

[Downloaded from vacres.pasteur.ac.ir on 2026-11-15]

[DOI: 10.61186/vacres.10.1.6]

RBD region, which is present in all coronavirus glycoproteins [13]. Most neutralizing antibodies against RBD are produced using the S1 subunit [14, 15]. Previous studies have shown that CD4+ responses to the S and N proteins which correlate with the neutralizing antibodies; especially, the role of S protein is crucial for the protective responses [16]. The M protein is the major structural protein of the virus coating that provides the overall form and is a type III membrane protein. The M protein is essential in the process of assembly, envelope formation, budding, and pathogenesis of the virus [17]. Another study has shown that M protein has been involved in escaping protective immune responses, specifically by interfering with IFNs signaling pathways [16].

Protein expression in prokaryotic systems is a cost-effective way for rapidly producing high amounts of a recombinant protein. The pET vector is designed for the expression of recombinant proteins in *Escherichia coli* which is controlled by a potent bacteriophage T7 promoter. In addition, the bacteria expression system can produce the desired product, which makes up more than half of the total cell protein after the induction [18]. In this study, we selected pET28a vector and *E. coli* (Rosetta) strain for the expression of a recombinant protein containing N, M and S1 proteins. This study provides a method for a rapid and cost-effective production of antigenic components of SARS-CoV-2 as recombinant proteins to be used in vaccine studies against COVID-19 or other future coronaviruses infections.

MATERIALS AND METHODS

Cloning and Extraction of Plasmid

The selected full-length sequences of N, M and S1 truncated of SARS-CoV-2 is described in our previous study [19]. The sequence was synthesized in a cloning vector (pcDNA3.1) between *Not*I (N-terminal) and *Nhe*I (C-terminal) restriction sites (Biomatic, Canada). The recombinant plasmid was transformed into *E. coli* (DH5α) strain. The bacteria colonies were selected on LB agar containing ampicillin. For plasmid extraction, a single colony was selected and inoculated into LB broth containing ampicillin for 16 h at 37°C. Mini-Kit (Favorgene, Taiwan) was used to extract plasmid from a single colony. After plasmid extraction, the plasmid and pET28a expression vector were digested by *Not*I and *Nhe*I restriction enzymes and purified from agarose gel with a gel-extraction kit (Favorgene, Taiwan). The restricted amplicons were ligated into pET28a using T4 DNA ligase. Following *E. coli* (Rosetta) transformation, selection on LB agar plate containing kanamycin, and plasmid preparation (as described above), the construct was then confirmed by PCR.

Expression of the Recombinant Protein

For the recombinant protein expression, a single recombinant colony was selected from the LB plate and inoculated into LB broth containing kanamycin for 16 h at 37°C. After 16 h , the bacterial culture was inoculated into 500 ml LB broth containing kanamycin, When OD_{600nm} reached 0.6, protein expression was induced by addition of 3 different concentrations of Isopropyl-D-thiogalactopyranoside (IPTG) (0.2, 0.5 and 1mM) for 4 deferent incubation time periods (2, 3, 4, and 16 h) at 37°C. The protein expression was confirmed by SDS-PAGE and Western blotting using an anti-His antibody (Abcam, USA).

Purification and Refolding of the Recombinant Protein

The recombinant protein was purified by Ni-NTA agarose (Qiagen, Germany). Briefly, the cell pellet was lysed by lysis buffer containing 8 M urea (100 mM NaH_2PO_4, 10 mM Tris-Cl, 8 M Urea) with pH 8, after 30 min incubation at room temperature (RT) and sonication, followed by adding Ni-NTA to the lysate and mixing it gently by shaking for 1 h at RT. After removing the bottom cap of the column, the flow-through collected. The column was washed twice with wash buffer (pH 6.3) and eluted twice by elution buffer (pH 4). The purified recombinant protein was refolded by urea buffer exchange. Briefly, the dialysis bag was pre-equilibrated with hot water for 5 min. The purified protein was poured into the bag, and the dialysis bag was placed in 5 M urea buffer for 3 h at 4°C. After 3 h, the buffer was changed with 2.5 M urea buffer for 3 h at 4°C. The buffer was then changed with PBS for 16 h at 4°C to completely remove the urea. The concentration of the recombinant protein was determined by NanoDrop 1000 spectrophotometer (NanoDrop Technologies, USA).

RESULTS

Plasmid Confirmation

For the plasmid confirmation, restriction enzymes (*Not*I and *Nhe*I) were used. The digestion products were recognized by 1% agarose gel as two bands of the insert (2743 bp) and pET28a expression vector (5369 bp). The digestion result showed that the plasmid was correctly constructed (Fig. 1B). The PCR results indicated that construct contains the insert (Fig. 1C).

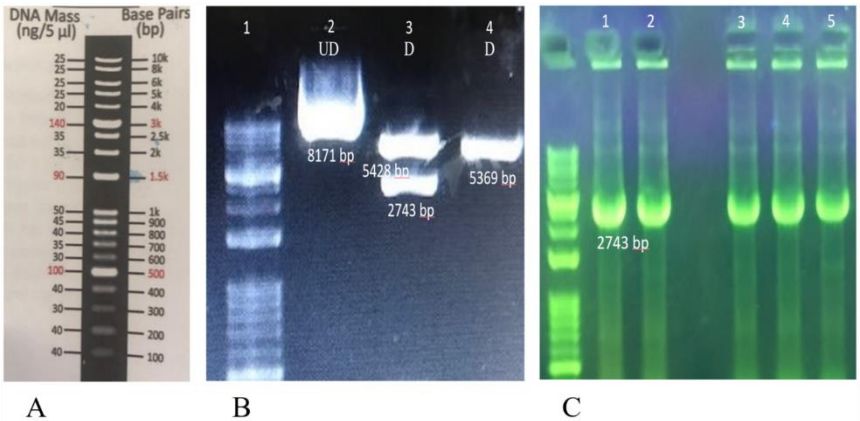

Fig. 1. Recombinant plasmid confirmation. **(A)** DNA marker 1kbp. **(B)** Digestion with restriction enzyme (*Not*I and *Nhe*I). Lane 1: DNA marker 1kbp (SMBIO). Lane 2: Undigested (8171 bp). Lane 3: Digested plasmid (2743 bp as gene) with *Not*I as N-terminal and *Nhe*I as C-terminal. Lane 4: Digested pET28a (5369 bp). **(C)** Recombinant plasmid identification by PCR. Lane 1-5: PCR product.

[Downloaded from vacres.pasteur.ac.ir on 2026-11-15]

[DOI: 10.61186/vacres.10.1.6]

Expression and Purification of the Recombinant Protein

Expression of the recombinant protein was performed in *E. coli* (Rosetta) and the expression optimization was evaluated in terms of IPTG concentrations (Fig. 2A), time of induction (Fig. 2B) and the selected colony (Fig. 2C). The best protein expression was obtained with 0.2 mM IPTG at 16 h . The recombinant protein was purified on a Ni-NTA column with the denaturing protocol (Fig. 2D, E). The molecular weight of the recombinant protein (94 kDa) and its histag was confirmed by SDS-PAGE gel and Western blotting, respectively (Fig. 2F).

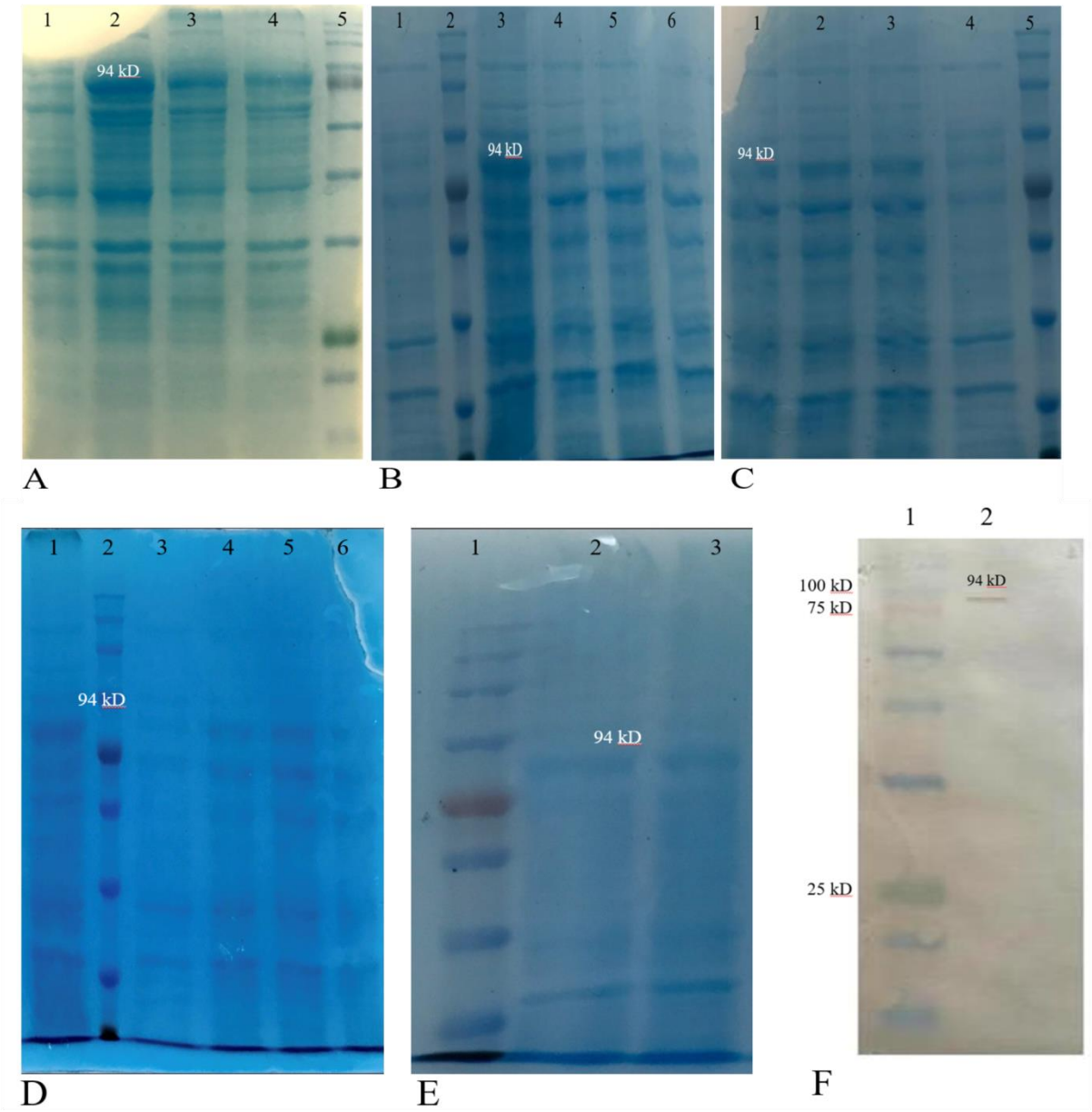

Fig. 2. The expression and purification of recombinant protein analyses. **(A)** IPTG concentration analyses of cell lysates of *E .coli* cells of recombinant protein after 16 h, Lane 1: Uninduced. Lane 2: 1mM IPTG. Lane 3: 0.5 mM IPTG. Lane 4: 0.2 mM IPTG. Lane 5: Protein marker (Sinaclon). **(B)** Time course of induction analyses, Lane 1: Uninduced. Lane 2: Protein marker. Lane 3: 16 h after induction. Lane 4: 4 h after induction. Lane 5: 3 h after induction. Lane 6: 2 h after induction. **(C)** Colony selection analyses, Lane 1: colony 1. Lane 2: colony 2. Lane 3: colony 3. Lane 4: Uninduced. Lane 5: Protein marker. **(D)** Purification process analyses. Lane 1: cell lysate of protein. Lane 2: Protein marker (Sinaclon). Lane 3: flow through. Lane 4-6: fraction of protein was washing 1-3. **(E)** Lane 1: Protein marker (Sinaclon). Lane 2, 3: fractions eluted. **(F)** Western blot analysis, Lane 1: Protein marker (Sinaclon). Lane 2: Purified recombinant protein.

DISCUSSION

The COVID-19 epidemic, caused by SARS-CoV-2, is a worldwide public health concern that needs immediate attention [20]. Identification and development of appropriate antigens is the first step for inducing the immune response by antibodies against SARS-CoV-2 and reliable development of vaccines [21, 22]. The previous results of the immune response to various SARS-CoV-2 proteins have shown that N protein is more

[Downloaded from vacres.pasteur.ac.ir on 2026-11-15]

[DOI: 10.61186/vacres.10.1.6]

immunogenic than other proteins [7]. Recent studies have indicated that immunogenicity by a full-length M protein lead to appropriate antibody induction in patients [23]. Other studies have shown that the interaction between the receptor and S1 proteins plays a major role in the infection and pathogenesis of SARS-CoV-2; therefore, S1 protein is a suitable therapeutic target protein or a vaccine component against COVID-19 [24]. Hence, we selected both full-length N, M proteins and truncated of S1 proteins in this study.

To develop subunit vaccines, it is important to select an optimal expression system that synthesizes the target protein while preserving its antigenic and immunogenic properties [25]. Prokaryotic protein expression systems are fast and cost-effective ways to produce the required amount of a recombinant protein [25]. In the current study, to evaluate the recombinant expression of our protein of interet, we first isolated the target protein gene from a pcDNA3.1 cloning vector and cloned it into pET28a expression vector. The expression of the 94 kDa recombinant protein was performed in *E. coli* (Rosetta) strain. The highest yield of the recombinant protein was obtained with 16 h of incubation with 0.2 mM of IPTG. Finally, the purified recombinant protein was confirmed by Western blotting. In this regard, we used pET28a expression vector because the pET family of vectors are good choices for gene expression in *E. coli* as they act under the tight control of T7 lac promoter and produce appropriate purity while saving production time and costs [26, 27].

Previously, Djukic *et al.* expressed a recombinant N protein in the bacterial expression system and showed that this protein induced IgG and IgM antibodies against SARS-CoV-2 [28]. The M protein of SARS-CoV-2 has strong antigenic properties and is less susceptible to mutation than protein S; hence, it is a good choice to be included as a component of a vaccine candidate against SARS-CoV-2 [29]. Moreover, M Protein has been shown to be as potent as S protein in blood samples from COVID-19 patients to stimulate and activate CD8 + cells [30]. Gao *et al.* have shown that the recombinant RBD protein expressed in *E. coli* (Rosetta) has biological activity and could be a good candidate for the development of new drugs against SARS-CoV-2 [27]. In a pre-clinical study, Yang *et al.* have shown that recombinant S protein is a suitable target for developing a safe and effective vaccine against SARS-CoV-2 infection [31]. In comparison with their study, we use all three N, M and S1 proteins in a construct to increase the possibility of the immunization. In conclusion, we have successfully obtained the expression and purification of a recombinant protein containing N, M and S fragments using an *E. coli* expression system, making it a viable and cost-effective option as a first step for the development of vaccines against SARS-CoV-2 or the future coronavirus infections upon further evaluations.

ACKNOWLEDGEMENTS

The project financially was financially supported by Pasteur Institute of Iran (Grant NO. 1159).

CONFLICT OF INTEREST

The authors declare they have no conflict of interests.

REFERENCES

1. Wang C, Horby PW, Hayden FG, Gao GF. A novel coronavirus outbreak of global health concern. The lancet. 2020;395(10223):470-3.
2. Phelan AL, Katz R, Gostin LO. The novel coronavirus originating in Wuhan, China: challenges for global health governance. Jama. 2020;323(8):709-10.
3. Liu D, An C, Bai Y, Li K, Liu J, Wang Q et al. A Novel Single-Stranded RNA-Based Adjuvant Improves the Immunogenicity of the SARS-CoV-2 Recombinant Protein Vaccine. Viruses. 2022;14(9):1854.
4. Poland GA, Ovsyannikova IG, Crooke SN, Kennedy RB, editors. SARS-CoV-2 vaccine development: current status. Mayo Clinic Proceedings; 2020: Elsevier.
5. Timani KA, Ye L, Ye L, Zhu Y, Wu Z, Gong Z. Cloning, sequencing, expression, and purification of SARS-associated coronavirus nucleocapsid protein for serodiagnosis of SARS. Journal of clinical virology. 2004;30(4):309-12.
6. Yang P, Wang X. COVID-19: a new challenge for human beings. Cellular & molecular immunology. 2020;17(5):555-7.
7. Smits VA, Hernández-Carralero E, Paz-Cabrera MC, Cabrera E, Hernández-Reyes Y, Hernández-Fernaud JR et al. The Nucleocapsid protein triggers the main humoral immune response in COVID-19 patients. Biochemical and biophysical research communications. 2021;543:45-9.
8. Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh C-L, Abiona O et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science. 2020;367(6483):1260-3.
9. Hoffmann M, Kleine-Weber H, Krüger N, Müller M, Drosten C, Pöhlmann S. The novel coronavirus 2019 (2019-nCoV) uses the SARS-coronavirus receptor ACE2 and the cellular protease TMPRSS2 for entry into target cells. BioRxiv. 2020.
10. Tortorici MA, Veesler D. Structural insights into coronavirus entry. Advances in virus research. Elsevier; 2019. p. 93-116.
11. Zhou G, Zhao Q. Perspectives on therapeutic neutralizing antibodies against the Novel Coronavirus SARS-CoV-2. International journal of biological sciences. 2020;16(10):1718.
12. Lan J, Ge J, Yu J, Shan S, Zhou H, Fan S et al. Crystal structure of the 2019-nCoV spike receptor-binding domain bound with the ACE2 receptor. BioRxiv. 2020.
13. Kirchdoerfer RN, Cottrell CA, Wang N, Pallesen J, Yassine HM, Turner HL et al. Pre-fusion structure of a human coronavirus spike protein. Nature. 2016;531(7592):118-21.
14. Shrwani K, Sharma R, Krishnan M, Jones T, Mayora-Neto M, Cantoni D et al. Detection of Serum Cross-Reactive Antibodies and Memory Response to SARS-CoV-2 in Prepandemic and Post–COVID-19 Convalescent Samples. The Journal of infectious diseases. 2021;224(8):1305-15.
15. Li R, Ma X, Deng J, Chen Q, Liu W, Peng Z et al. Differential efficiencies to neutralize the novel mutants B. 1.1. 7 and 501Y. V2 by collected sera from convalescent COVID-19 patients and RBD nanoparticle-vaccinated rhesus macaques. Cellular & molecular immunology. 2021;18(4):1058-60.
16. Gazzinelli-Guimaraes PH, Sanku G, Sette A, Weiskopf D, Schaughency P, Lack J et al. Antigenic determinants of SARS-CoV-2-specific CD4+ T cell lines reveals M protein-driven dysregulation of interferon signaling. medRxiv. 2022.
17. Kota D. Coronavirus and its structure. International Journal of Collaborative Research on Internal Medicine & Public Health. 2021;13(6):1-2.
18. Shilling PJ, Mirzadeh K, Cumming AJ, Widesheim M, Köck Z, Daley DO. Improved designs for pET expression plasmids increase protein production yield in Escherichia coli. Communications biology. 2020;3(1):1-8.
19. Sadat SM, Aghadadeghi MR, Yousefi M, Khodaei A, Sadat Larijani M, Bahramali G. Bioinformatics analysis of SARS-CoV-2 to approach an effective vaccine candidate against COVID-19. Molecular Biotechnology. 2021;63(5):389-409.
20. Behzadifar M, Imani-Nasab M-H, Martini M, Ghanbari MK, Bakhtiari A, Bragazzi NL. Universal Health Coverage to counteract the economic impact of the COVID-19 infection: Current practices and ethical challenges. Journal of Preventive Medicine and Hygiene. 2020;61(4):E520.
21. Siu Y, Teoh K, Lo J, Chan C, Kien F, Escriou N et al. The M, E, and N structural proteins of the severe acute respiratory syndrome coronavirus are required for efficient assembly, trafficking, and release of virus-like particles. Journal of virology. 2008;82(22):11318-30.
22. Premkumar L, Segovia-Chumbez B, Jadi R, Martinez DR, Raut R, Markmann AJ et al. The receptor-binding domain of the viral spike protein is an immunodominant and highly specific target of antibodies in SARS-CoV-2 patients. Science immunology. 2020;5(48):eabc8413.
23. Lin L, Ting S, Yufei H, Wendong L, Yubo F, Jing Z. Epitope-based peptide vaccines predicted against novel coronavirus disease caused by SARS-CoV-2. Virus research. 2020;288:198082.

[Downloaded from vacres.pasteur.ac.ir on 2026-11-15]

[DOI: 10.61186/vacres.10.1.6]

24. Veljkovic V, Vergara-Alert J, Segalés J, Paessler S. Use of the informational spectrum methodology for rapid biological analysis of the novel coronavirus 2019-nCoV: prediction of potential receptor, natural reservoir, tropism and therapeutic/vaccine target. F1000Research. 2020;9.
25. Bobbala S, Hook S. Is there an optimal formulation and delivery strategy for subunit vaccines? Pharmaceutical research. 2016;33(9):2078-97.
26. Dobrijevic D, Nematollahi LA, Hailes HC, Ward JM. pET expression vector customized for efficient seamless cloning. Biotechniques. 2020;69(5):384-7.
27. Gao X, Peng S, Mei S, Liang K, Khan MSI, Vong EG et al. Expression and functional identification of recombinant SARS-CoV-2 receptor binding domain (RBD) from E. coli system. Preparative Biochemistry & Biotechnology. 2022;52(3):318-24.
28. Djukic T, Mladenovic M, Stanic-Vucinic D, Radosavljevic J, Smiljanic K, Sabljic L et al. Expression, purification and immunological characterization of recombinant nucleocapsid protein fragment from SARS-CoV-2. Virology. 2021;557:15-22.
29. Lopandić Z, Protić-Rosić I, Todorović A, Glamočlija S, Gnjatović M, Ćujic D et al. IgM and IgG immunoreactivity of SARS-CoV-2 recombinant M protein. International Journal of Molecular Sciences. 2021;22(9):4951.
30. Grifoni A, Weiskopf D, Ramirez SI, Mateus J, Dan JM, Moderbacher CR et al. Targets of T cell responses to SARS-CoV-2 coronavirus in humans with COVID-19 disease and unexposed individuals. Cell. 2020;181(7):1489-501. e15.
31. Yang J, Wang W, Chen Z, Lu S, Yang F, Bi Z et al. A vaccine targeting the RBD of the S protein of SARS-CoV-2 induces protective immunity. Nature. 2020;586(7830):572-7.

[Downloaded from vacres.pasteur.ac.ir on 2026-11-15]

[DOI: 10.61186/vacres.10.1.6]