BIOSCIENCES BIOTECHNOLOGY RESEARCH ASIA, June 2024. Vol. **21**(2), p. 689-703

Recombinant Expression and Purification of Novel COVID-19 Vaccine Candidate in *Escherichia coli* BL21 (DE3)

Alok Singh[1]*, Navneet Verma[2], Prevesh Kumar[2], Diksha[2] and Iqra Hasan[3]

[1]Research Scholar, Faculty of Pharmacy, IFTM University, Moradabad, Uttar Pradesh, India.
[2]Pharmacy Academy, IFTM University, Moradabad, Uttar Pradesh, India.
[3]School of Biotechnology, IFTM University, Moradabad, Uttar Pradesh, India.

http://dx.doi.org/10.13005/bbra/3256

(Received: 12 February 2024; accepted: 20 May 2024)

COVID-19, the global pandemic, infected and killed many human beings across the world. The sudden onset and global spread of the disease necessitated the development of an efficient vaccine for mass vaccination. The present study provides the data for the expression and purification of a vaccine candidate against the SARS-CoV2 virus. The beauty of this vaccine is the employment of multiple epitopes targeting the structural and non-structural proteins of the virus, thus inhibiting the viral infection and replication. The study data showed that the recombinant vaccine candidate was sequestered into inclusion bodies in Escherichia coli (E. coli) BL21 (DE3). In order to maximize protein recovery, protein solubilization and refolding was optimized using mild chaotropic agents. Further, anion exchange (AEX) chromatography was used as a negative chromatography to remove other protein impurities and recover the protein of interest in the flow-through. The cation exchange (CEX) chromatography step provided pure protein, but the protein recovery was reduced. The final purified protein showed the presence of NSP9 and RBD when probed with antibodies against these epitopes. The study demonstrated that a multiple epitope vaccine can be successfully expressed using E. coli BL21 (DE3) as the host. However, further studies are required to prove the efficacy of the vaccine candidate.

Keywords: Chromatography; COVID-19; Expression; Purification; Refolding; Vaccine.

Immunization is the foundation of public health for any nation. The first vaccination was done by Edward Jenner against small pox in a 8 year old boy using cowpox lesion scratching from a milk maid[1–3]. Different types of vaccines are available against different disease-causing pathogens. Vaccines are usually classified into live attenuated and killed inactivated vaccines. Live attenuated vaccines have weakened but viable pathogen as antigens, while in killed vaccines; pathogens are inactivated by heat or chemicals. Over years, vaccines have been developed using specific biomolecules like proteins, DNA and mRNA[4–6]. Subunit vaccine makes use of the highly antigenic portion of viral and bacterial proteins, while conjugate vaccines are bacterial lipopolysaccharides covalently bound to an antigenic carrier protein to provide broad spectrum immunity. The protein-based vaccines can be produced using recombinant DNA technology. DNA and mRNA based vaccines have recently become strong candidates[5,6].

Recent pandemic of Coronavirus Disease 2019 (COVID-19) witnessed fast track research for

*Corresponding author E-mail: alokthevictor@gmail.com

developing a vaccine against its causative agent, severe acute respiratory syndrome coronavirus-2 (SARS CoV-2). Since 2019, more than 760 million people were infected and the virus killed 6.9 million deaths globally (https://www.who.int). SARS-CoV2 is a positive single-stranded ribonucleic acid (RNA) virus with a genome size of 26-36 kilobases and belongs to the *Coroniviridae* family. Several studies have shown that SARS-CoV2 has genetic similarities to the earlier known coronaviruses, severe acute respiratory syndrome coronavirus(SARS-CoV)and Middle East respiratory syndrome coronavirus (MERS-CoV), which also infect humans[7–9].The genome of SARS-CoV2 encodes four main structural proteins, i.e., spike glycoprotein (S), membrane protein (M), envelope (E) and nucleocapsid (N) proteins. These structural proteins are essential to maintain function and structure of the virus. The viral genome also carries genes for non-structural proteins (NSP) and other accessory proteins[8,9]. Sixteen NSPs are expressed from the 5'-end of the RNA genome and each NSP has been shown to have a significant function in establishing infection and immune evasion[10]. NSP7, NSP8 and NSP12 (also known as RNA-dependent RNA polymerase [RdRP]) form the replication-transcription complex (RTC) that is important for viral replication and transcription[11]. NSP9 is a RNA-binding protein shown to have a role in the formation of the RTC[12,13]. Zong *et al.* (2023) have revealed that NSP8 can induce mitophagy through mitochondrial damage[14]. A recent study has shown that targeting NSP9 can hinder its binding to NSP12 and thus curbing viral replication. Further, NSP9 is well conserved among coronaviruses, with a sequence identity of 98% between SARS-CoV and SARS-CoV2[15].

Different approaches for targeting the virus have been employed while developing vaccines for COVID-19. The protein subunit vaccines mainly targeted either the receptor biding domain (RBD) or S1 subunit of the spike protein of SARS-CoV2. Other vaccines were based on novel approaches of delivering SARS-CoV2 antigen by injecting antigen expression mRNA or virus-like particles[16]. An inactivated whole-virion vaccine was developed in India and was demonstrated to have good efficacy for mass immunization[17].Owing to the significance of NSPs in viral replication and establishment of infection, they can also act as important vaccine targets. In accordance to these observations, the study provides detailed information regarding the expression of a novel polymeric epitope vaccine candidate using RBD (a structural protein) and NSP9 (a non-structural protein) of SARS-CoV2 in *E. coli* BL21 (DE3). Furthermore, the research presents data about its purification using protein refolding, AEX and CEX chromatography from inclusion bodies and characterization by western blotting using polyclonal antibodies.

MATERIALS AND METHODS

Optimization of Expression conditions

The glycerol stock of *E. coli*BL21 (DE3) strain carrying the expression construct used in the study was obtained from Aurigene Pharmaceutical Services Limited (APSL), Hyderabad. The stock was streaked on a Luria-Bertani (LB) agar plate to obtain isolated colonies. Kanamycin (50 µg/mL) was used as the selection pressure in all the culturing steps, unless mentioned otherwise. Single colony was grown in 5 mL kanamycin containing LB broth by incubating at 37ºC, 200 rpm for 16-18 h. The overnight grown culture was used to inoculate 8 different flasks (F1-F8), each containing 30 mL LB broth + kanamycin for optimizing time, temperature and isopropyl-â-D-1-thiogalactopyranoside (IPTG) concentrations. All the flasks were incubated at 37ºC, 200 rpm until the $O.D._{600}$ reached 0.65-0.75. Table 1 shows the scheme of incubation conditions and IPTG addition to each flask to determine the optimal condition for expression.

At the end of induction period, $O.D._{600}$ was checked and 2.0 $O.D._{600}$ equivalent cells from all the flasks were sampled for expression analysis using sodium dodecyl sulfate – polyacrylamide gel electrophoresis (SDS-PAGE) as described earlier[18,19]. Briefly, required volume of culture was diluted to 500 µL using 250 mM Tris buffer, pH 7.5 and equal volume of 2X reducing Laemmli buffer (Bio-Rad) was added. The samples were heated at 90ºC for 20 min, and centrifuged at 9023 ×g, 4ºC for 5 min. Total cell lysate (15 µL) was loaded onto 12% pre-casted SDS-PAGE gel (Bio-Rad) with 4% stacking gel for determination of expression. Protein bands were visualized using Coomassie Brilliant Blue. The gel was incubated

in gel staining solution (1 g of Coomassie Brilliant Blue in 1 L of mix of50% [v/v] methanol, 10% [v/v] glacial acetic acid and 40% [v/v] water for injection [WFI]) for 20 min at room temperature. Extra stain was removed incubating the gel in destaining solution(7% [v/v] glacial acetic acid, 81% [v/v] WFI and 12% [v/v] methanol) for 20 min at room temperature[20,21].

In order to check for consistency of expression, six replicate flasks (F_I1 to F_I6) were cultured and induced with the finalized expression conditions, i.e., 37°C, 4 h and induction with 2 mM IPTG as described above. Other six flasks were cultured for un-induced controls ($F_{UI}1$ to $F_{UI}6$).

Localization of expressed protein

Once the expression conditions were optimized, 500 mL culture was grown as described earlier and induced for protein expression. Culture was homogenized and used for determining the localization of expressed recombinant protein as described elsewhere[22]. Briefly, 10 mL of culture was harvested by centrifuging at 10000 ×g, 4ºC for 15 min. 1 mL of cell lysis buffer (50 mM Tris-HCl, 5 mM ethylene diamine tetra acetic acid [EDTA], 1 mM protease inhibitor, phenylmethylsulfonyl fluoride [PMSF], pH 7.0 and conductivity5 mS/cm) was utilized to resuspension the cell culture pellet, with the help of Polytron homogenizer. Further, lysis was done using Panda plus using 12000 bar pressure, at 4ºC. After 5 passes, lysate was centrifuged at 16500 ×g, 4ºC for 60 min. The supernatant was collected in a different tube. The complete cell lysate, supernatant and cell debris pellet were analyzed on SDS-PAGE for localization of protein using 20 µL of samples as described earlier.

Large scale culture growth by Fed-batch Fermentation

Pre-seed culture was grown by inoculating with glycerol stock in 100 mL LB broth containing kanamycin. The flask was incubated at 37ºC and 200 rpm for 16–18 h.

After overnight growth, the $O.D._{600}$was measured and at an $O.D._{600}$ of about 4.2-5.0, seed flask for fermentor was inoculated. For seed flask, 40-60 mL pre-seed culture was added into kanamycin containing LB broth (200 mL). The flask was incubated at 37ºC, 200 rpm for 2-3 h until the $O.D._{600}$ reached >2.5. The composition of the fermentor media was as described in Table 2.

Struktol (5%) was used as anti-foaming agent. Kanamycin (75 µg/mL) was added before seeding of fermentor and also 1 h before induction with 2 mM IPTG. Feed medium was composed of 60% [w/v] glucose, 40% [w/v] yeast extract and 2.5g magnesium sulfate per 100 g of glucose. Fermentation parameters were set as follows: incubation temperature = 37°C, pH = 6.9, dissolved oxygen (DO) =40%, 200rpm, air = 0.5 lpm and overlay air = 0.5 lpm.

Induction was started once the $O.D._{600}$ reached 80-90 and was carried out for 4-5 h. Sampling was done every hour for measuring $O.D._{600}$, pH and glucose levels. Once induction was complete, the temperature of fermentor was reduced to 8-10°C before harvesting the culture. The culture was harvested by centrifuging at 18000 ×g, 4-6°C for 30 min. The obtained cell pellet was kept at -20°C for storage.

Isolation of inclusion bodies

Inclusion bodies were isolated as described by Eggenreich *et al.* (2020) with some modifications[23]. Briefly, resuspension of 2.5 g of cell pellet was done using a Polytron homogenizer in 50 mL cold lysis buffer (50 mM Tris-HCl, 5 mM EDTA, 1 mM PMSF at pH 7.0 and conductivity 5 mS/cm). Further, the cell lysate was prepared by passing through Panda plus homogenizer five times at 12000 bar pressure and 4°C. The sample was clarified using centrifugation at 16500 ×g, 4°C for 60 min. The wet weight of cell pellet was measured and resuspended in 50 mL cold wash buffer (50 mM Tris-HCl, 500 mM NaCl, and 1% [v/v] Triton X-100, at pH 7.0 and conductivity 48 mS/cm). Centrifugation was done at 16500 ×g, 4°C for 60 min. The washing step was repeated again and the wet weight of resultant pellet was measured and resuspended in 50 mL cold WFI to remove detergents and salts from the inclusion bodies. After centrifuging the mixture at 16500 ×g, 4°C for 60 min, the pellet, which consisted of inclusion bodies, was weighed. SDS-PAGE analysis was performed after every centrifugation step as described earlier.

Solubilization and refolding of protein

Solubilization and refolding of protein from inclusion bodies was carried out using the protocol described in earlier studies[24–26]. The pure inclusion bodies obtained after WFI wash were dissolved in 10 mL WFI and 40 mL solubilization buffer (2 M urea, 50 mM Tris and 1

mM dithiothreitol [DTT], pH 12.0) by continuously stirring using a magnetic stirrer for 30 min at 4°C. Centrifugation of solubilized inclusion bodies at 14000 ×g, 4°Cfor 60 min was done to remove any insoluble aggregates.

Protein refolding was done by diluting the filtered supernatant 10 times with refolding buffer (0.2 M urea, 50 mM Tris, 10%[w/v] sucrose, 10 mM EDTA and 250 mM arginine, pH 7.2, conductivity 18 mS/cm) and continuously mixing using magnetic stirrer for 12-16 h at 4°C. After 15 h, the aggregates and unfolded proteins were removed from the refolded protein by centrifugation at 14000 ×g, 4°C for 30 min. The pH and conductivity were measured for the refolded protein sample. SDS PAGE analysis was performed for both supernatant and pellet after solubilization and refolding processes. Based on SDS PAGE data, the supernatant was filtered through 0.45 ìm filter and used for purification. Protein content was determined using the Pierce™ BCA protein assay kit (Thermo Scientific).

Anion Exchange (AEX) Chromatography

The refolded protein was purified using AEX column in the flow-through mode using AKTA Pure chromatography system (Cytiva). The strong anion exchanger resin, Q-XL Sepharose (Cytiva; 5 mL), was packed in XK 16 column. Column was equilibrated with 10 column volume (CV), i.e. 50 mL of equilibration buffer (0.2 M Urea, 50 mM Tris, 10%[w/v] sucrose, 10 mM EDTA and 250 mM arginine, pH 7.2, conductivity 18 mS/cm), and then the refolded protein sample was loaded onto the column. Post load washing was done using 5 CV equilibration buffer, followed by elution with 5 CV of elution buffer (0.2 M Urea, 50 mM Tris, 10%[w/v] sucrose, 10 mM EDTA and 250 mM arginine, and 1 M NaCl, pH 7.2, conductivity 78 mS/cm). SDS-PAGE analysis was done to analyze the purity levels as described earlier.

Concentration and Buffer exchange by Tangential Flow Filtration (TFF)

Since the protein was obtained in the flow-through of AEX chromatography, the protein sample concentration and buffer exchange were done using a 100 kilo Daltons (kDa) molecular weight cut-off TFF cassette (low protein binding polyethersulfone [PES] membrane). The protein sample was concentrated five folds, by reducing the volume from 500 mL to ~100 mL based on the $O.D._{280}$ values not exceeding 0.5 mg/mL. The transmembrane pressure (TMP) during ultrafiltration was kept constant at 0.25 Bar. Further, the diafiltration buffer (50 mM Tris, 10% [w/v] sucrose, 0.01% [v/v] Pluronic F68, and 250 mM arginine, pH 7.2, conductivity 5-7 mS/cm,) was substituted for the original buffer.

After completion of diafiltration, TFF retentate (80 mL) was stored in a bottle. The cassette was flushed twice with diafiltration buffer (10 mL each) to recover protein remaining in the cassette. The TFF retentate was filtered using 0.45 µm filter. SDS-PAGE analysis was done to analyze the purity levels.

Cation exchange (CEX) chromatography

The TFF retentate was purified using CEX chromatography to remove impurities using SP Sepharose XL resin from Cytiva and AKTA Pure chromatography system (Cytiva). Resin (5 mL) was packed in XK 16 column. Column was equilibrated with 10 column volume (CV), i.e. 50 mL of equilibration buffer (50 mM Tris, 10% sucrose, 0.01% Pluronic F68 and 250 mM arginine, pH 7.2, conductivity 5-7 mS/cm). 100 mL TFF retentate was loaded onto the column and post load washing was done using 5 CV of equilibration buffer. 5 CV elution buffer (50 mM Tris, 10% [w/v] sucrose, 0.01% [v/v] Pluronic F68, 250 mM arginine and 500 mM NaCl, pH 7.2, conductivity43 - 46 mS/cm) was applied for elution of the protein.

Elution fractions (7 fractions of 3 mL each) were collected based on $O.D._{280}$ values. SDS-PAGE analysis was done to analyze the purity levels.

Ultrafiltration and diafiltration (UFDF)

Pooled CEX elution fraction (21 mL) was used for this step. A 100 kDa molecular weight cut-off Amicon device was used. Concentration was performed by centrifugation at 2240 ×g, 4°C for 3-5min. The protein was concentrated to a concentration not more than 0.5 mg/mL. Further the buffer was exchanged with 50 mM Tris, 10% [w/v] sucrose, 0.01% [v/v] Pluronic F68 and 250 mM arginine, pH 7.2, conductivity 6 mS/cm. After buffer exchange, pH and conductivity of the protein sample was measured and was observed to match with diafiltration buffer. The final volume of UFDF retentate was 5 mL. The sample was filtered through 0.45 µm filter, followed by sterile filtration with a 0.2 µm PES membrane syringe filter.

Western blotting

Final purified protein (loading amount of 3 µg and 5 µg) was resolved on 12% SDS-PAGE gel with molecular weight marker (15 kDa – 180 kDa) in duplicates. The gel was briefly washed with water and then protein was transferred onto activated polyvinylidene fluoride [PVDF] membrane. The transfer was done using 1X Tris-Glycine-SDS-Methanol buffer at a constant voltage of 30V for 16 h at 2-8ºC. After completion of transfer, the membrane was blocked with 2.5% [w/v] non-fat milk powder in1X phosphate buffered saline (37 mM Sodium chloride [NaCl], 2.7 mM Potassium chloride [KCl], 4.3 mM Na_2HPO_4, 1.47 mM KH_2PO_4, pH 7.4) with 0.05% [v/v] Tween 20 (1X PBST) for 30 min with gentle rocking. Once blocked, 5 mL 1X PBST was used to wash the membrane thrice for 5 min each on a rocker. The membrane was cut into two halves and incubated with two different primary antibodies (Anti-SARS-CoV-2 NSP-9 polyclonal antibody [In house, APSL], dilution 1: 1500 in 1X PBST and Anti-SARS-CoV-2 RBD polyclonal antibody [Gene Tex, Cat# GTX635692], dilution 1:1500 in 1X PBST) for 2 h at room temperature (RT) with gentle rocking. Excess and unbound primary antibodies were washed off by washing thrice with 1X PBST for 5 min each on rocker. Further, the membranes were probed with corresponding secondary antibodies (for NSP-9: Goat Anti-Mouse IgG HRP [Invitrogen, Cat# A16066], 1:5000 in 1XPBST and for RBD: Goat Anti- Rabbit IgG HRP [Dako, Cat# P0448], 1:5000 in 1X PBST) for 1 h at RT with gentle rocking. This was followed by washing thrice with 1X PBST for 5 min each on rocker to remove excess secondary antibody. The membranes were finally washed twice with 1X PBS for 5 min each on rocker and then reaction was developed with 3,32 -Diaminobenzidine (DAB, 0.05% [w/v] in 1X PBST) and H_2O_2(0.5% [v/v]) in dark. Once the expected bands appeared on the blot, the developing solution was drained; the membrane was washed in purified water and dried before scanning the blots.

Table 1. Conditions tested for optimizing expression of protein

Flask ID	Induction Temperature	Induction time period	IPTG concentration
F1	37ºC	4 h	0 mM
F2	37ºC	4 h	0.5 mM
F3	37ºC	4 h	1.0 mM
F4	37ºC	4 h	2.0 mM
F5	16ºC	16-18 h	0 mM
F6	16ºC	16-18 h	0.5 mM
F7	16ºC	16-18 h	1.0 mM
F8	16ºC	16-18 h	2.0 mM

Table 2. Composition of fermentor medium

Chemical Composition	Concentration (g/L)
Potassium dihydrogen phosphate (KH_2PO_4)	4.0
Di-potassium hydrogen phosphate (K_2HPO_4)	4.0
Di-sodium hydrogen phosphate dodecahydrate ($Na_2HPO_4.12H_2O$)	7.0
Ammonium sulfate ($[NH_4]_2SO_4$)	1.2
Ammonium chloride (NH_4Cl)	0.2
Magnesium sulfate heptahydrate ($MgSO_4.7H_2O$)	2.4
Yeast extract	10.0
Glucose	10.0
Trace elements	
Ferrous sulfate heptahydrate ($FeSO_4.7H_2O$)	40.0 mg
Calcium chloride ($CaC1_2.2H_2O$)	40.0 mg
Manganese(II) sulfate monohydrate ($MnSO_4.H_2O$)	10.0 mg
aluminum chloride hexahydrate ($AlCl_3.6H_2O$)	10.0mg
Cobalt (II) chloride hexahydrate ($CoC1_2.6H_2O$)	4.0 mg
Zinc sulfate heptahydrate ($ZnSO_4.7H_2O$)	2.0mg
Sodium molybdate dihydrate ($Na_2MoO_4.2H_2O$)	2.0mg
Copper(II) chloride dihydrate ($CuC1_2.2H_2O$)	1.0mg
Boric acid (H_3BO_3)	0.5mg

RESULTS

Optimization of expression

Different time-temperature and IPTG concentration combinations were used to analyze the level of protein expression as mentioned in materials and methods. Protein expression was analyzed by running crude cell lysates on SDS-PAGE. While the $O.D._{600}$ increased post induction, highest $O.D._{600}$ was observed in cultures induced with 2 mM IPTG (Table 3).

It was observed that induction with 2 mM IPTG resulted in highest expression (Fig. 1, Lane 4 and Lane 7) as compared to 1 mM (Fig. 1, Lanes 5 and 8) and 0.5 mM IPTG (Fig. 1, Lanes 6 and 9). Preliminary experiments showed that the extent of expression induction did not change with the change in temperature or time duration. Thus, the conditions of expression were finalized to be 37°C, 4 h and induction with 2 mM IPTG.

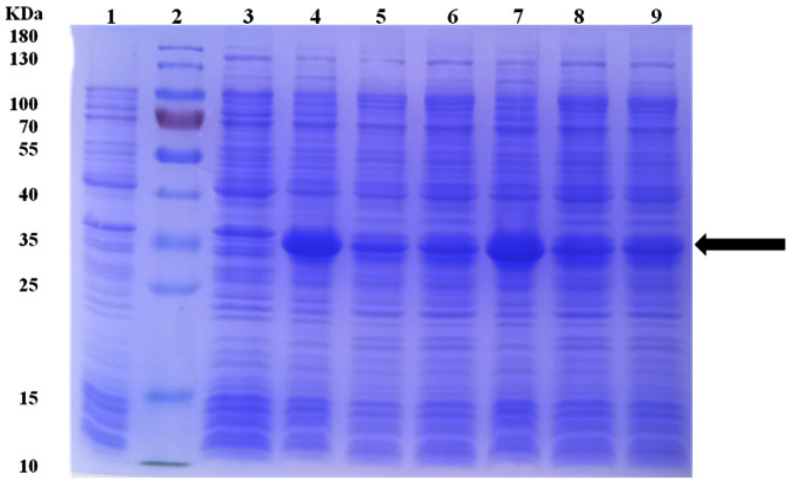
Further, six replicate flasks were cultured to check for consistency. All the six flasks showed similar $O.D._{600}$ at the time of harvest (Table 4). SDS-PAGE analysis of crude cell lysates showed similar levels of expression of the protein of interest (Fig. 2 A and B).

Localization of expressed protein

Upon confirmation of culture and induction conditions, localization of expressed protein was determined. The protein was expressed

Table 3. Culture conditions and $O.D._{600}$ of cultures at the time of harvest after induction in expression analysis

Flask no.	Antibiotic	Temp (°C)	Rotation speed (rpm)	$O.D._{600}$ at Induction	IPTG Conc. (mM)	Induction time (h)	Induction temperature (°C)	$O.D._{600}$ at harvest
1	Kanamycin	37±1	200	0.72	0	4	37±1	8.16
2	(50 µg/mL)			0.72	0.5	4	37±1	6.65
3				0.71	1.0	4	37±1	6.64
4				0.72	2.0	4	37±1	14.97
5				0.73	0	16-18	16±1	8.04
6				0.71	0.5	16-18	16±1	7.25
7				0.72	1.0	16-18	16±1	8.16
8				0.75	2.0	16-18	16±1	14.90



Lane 1: Culture lysate (un-induced, 37°C, 4 h); Lane 2: Molecular weight marker (10 kDa – 180 kDa); Lane 3: Culture lysates (un-induced, 16°C, 16-18 h); Lane 4: Culture lysate (2 mM IPTG, 37°C, 4 h); Lane 5: Culture lysate (1 mM IPTG, 37°C, 4 h); Lane 6: Culture lysate (0.5 mM IPTG, 37°C, 4 h); Lane 7: Culture lysate (2 mM IPTG, 16°C, 16-18 h); Lane 8: Culture lysate (1 mM IPTG, 16°C, 16-18 h); and Lane 9: Culture lysate (0.5 mM IPTG, 16°C, 16-18 h).

Fig. 1. SDS-PAGE gel showing expression of protein at different incubation and induction conditions

as inclusion bodies and not as soluble, cytoplasmic fraction (Fig. 3).

Once it was deduced that the protein was expressed and sequestered into inclusion bodies, fed-batch fermentation was carried out for a large-scale batch to optimize purification of the protein. The details of parameters (glucose levels, pH and $O.D._{600}$) measured over fermentation process are provided in Table 5. The values in bold indicate the time of induction of protein expression.

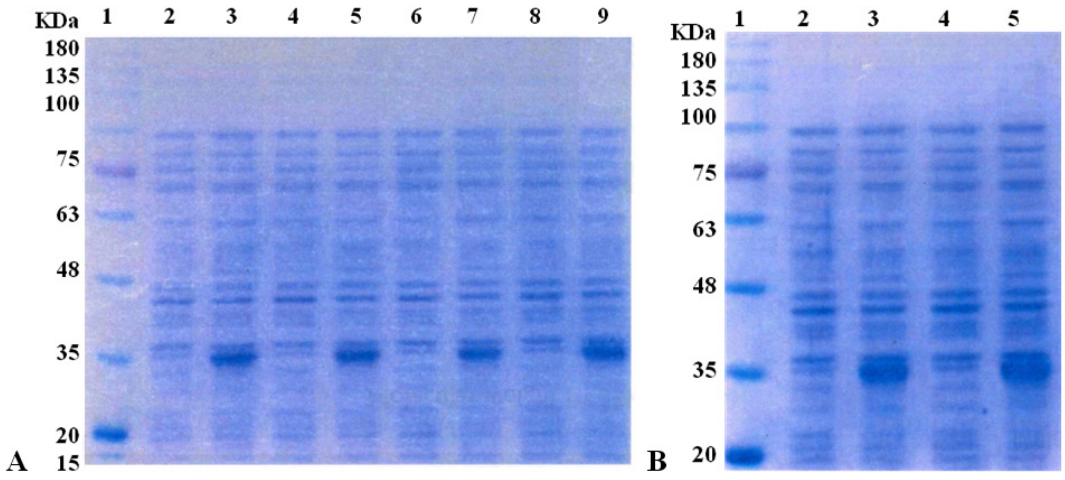
Inclusion bodies isolation, protein solubilization and refolding

The process was started with 2.5 g of cell pellet and the weight of inclusion bodies achieved after the final wash with WFI was 0.22 g. Figure 4A shows that pure protein is obtained after every washing step during isolation of inclusion bodies.

Although purification of expressed protein from inclusion bodies becomes easier, the refolding process can decrease the yield of active protein of interest[27]. To this end, BCA protein assay was used to determine protein content after every step of solubilization and refolding and percentage step recovery was calculated. It was observed that using mild chaotropic agent like urea to solubilize inclusion bodies and refold protein facilitated achievement of partially purified protein (Figure 4B). The percentage step recovery of protein after refolding was 90.76% as observed by protein estimated using BCA protein assay indicating minimal loss of protein (Table 6).

Table 4. Culture conditions and $O.D._{600}$ of cultures at the time of harvest for replicate flasks

Flask	Antibiotic	Temp (°C)	Rotation speed (rpm)	$O.D._{600}$ at Induction	IPTG Conc. (mM)	Induction time (h)	$O.D._{600}$ at harvest
F_I1	Kanamycin	37±1	200	0.72	2	4	14.25
F_I2	(50 µg/mL)			0.72	2		14.39
F_I3				0.71	2		17.98
F_I4				0.72	2		14.58
F_I5				0.73	2		15.01
F_I6				0.74	2		15.20



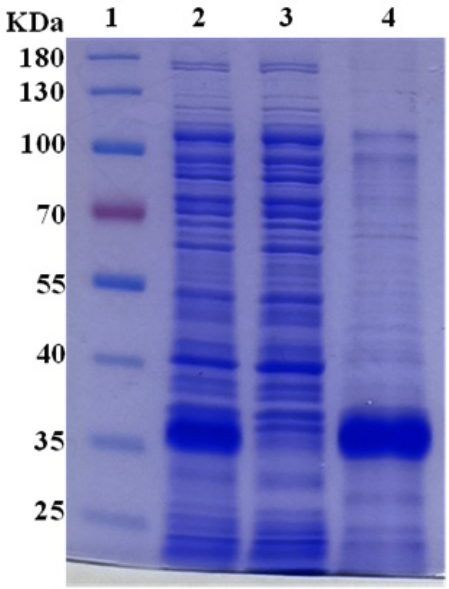
Whole cell lysates (15 ìL) were loaded for each sample **(A)** Lane 1: Molecular weight marker (15 kDa – 180 kDa); Lane 2: Un-induced flask $F_{UI}1$; Lane 3: Induced flask F_I1; Lane 4: Un-induced flask $F_{UI}2$; Lane 5: Induced flask F_I2; Lane 6: Un-induced flask $F_{UI}3$; Lane 7: Induced flask F_I3; Lane 8: Un-induced flask $F_{UI}4$; Lane 9: Induced flask F_I4. **(B)** Lane 1: Molecular weight marker (15 kDa – 180 kDa), Lane 2: Un-induced flask $F_{UI}5$; Lane 3: Induced flask F_I15 Lane 4: Un-induced flask $F_{UI}6$; Lane 5: Induced flask F_I6.

Fig. 2. SDS-PAGE analysis for consistency of protein expression in replicate flasks

Purification of expressed protein using AEX and concentration of flow-through by TFF

AEX chromatography step acted as a negative purification to remove extraneous protein impurities. Since theoretical isoelectric pH (pI) of the protein of interest was 8.32 and a pH of 7.2 was maintained throughout the AEX process, the protein remained positively charged and hence did not bind the resin and was present in the flow-through. Other protein impurities were eluted later (Figure 5A).

As refolded protein from inclusion bodies was used as loading material, the load and flow-through did not show much impurity for the volume (20 µL) loaded onto the gel (Fig. 5A Lanes 2 and 3). However, the elute fractions were more concentrated and hence showed the presence of impurities when same volume (20 µL) was loaded onto the gel (Fig. 5A, Lanes 4 to 10). The chromatogram of the AEX chromatography run provided the profiles of absorbance at 280 nm (A280), conductivity, absorbance at 260 nm (A260) and pH (Figure 5C).



Lane 1: Molecular weight marker (25 kDa – 180 kDa); Lane 2: Whole cell lysate; Lane 3: Supernatant; Lane 4: Cell pellet.

Fig. 3. SDS-PAGE gel depicting localization of protein of interest in inclusion bodies

Since the volume of flow-through was high (500 mL), it was concentrated using TFF. Concentration helps in preparing the sample for the next steps of purification and also allows exchange of buffers to the more compatible ones for protein stability and further purification methods. Figure 5B exhibited the SDS-PAGE analysis of the UFDF permeate and retentate. Further, protein estimation by BCA assay showed that AEX retained 96.68% of the protein in the flow-through and post TFF, the retentate showed a recovery of 99.22% (Table 6).

Purification of expressed protein using CEX and concentration of pooled elute fractions by UFDF

Since the purified protein showed some high and low molecular weight impurities after AEX and TFF, CEX using SP Sepharose XL resin was employed to further purify the protein. Figure 6A showed the SDS-PAGE profile of elute fractions obtained after CEX chromatography. The protein elutes appeared to be pure and devoid of impurities. The chromatogram for the CEX chromatography run was presented in Figure 6C and 6D. However, the protein recovery from this step was only 11.41% (Table 7).

The elute fractions of CEX were pooled and concentrated to ~5 mL using an amicon centrifugal device. UFDF output was the final product (Fig. 6B) and was used for analyzing the correctness of epitopes by western blotting.

Detection of epitopes by Western blotting

Since the vaccine candidate carries multiple epitopes like RBD and NSP9 of SARS-CoV2 virus, the purified protein was probed with

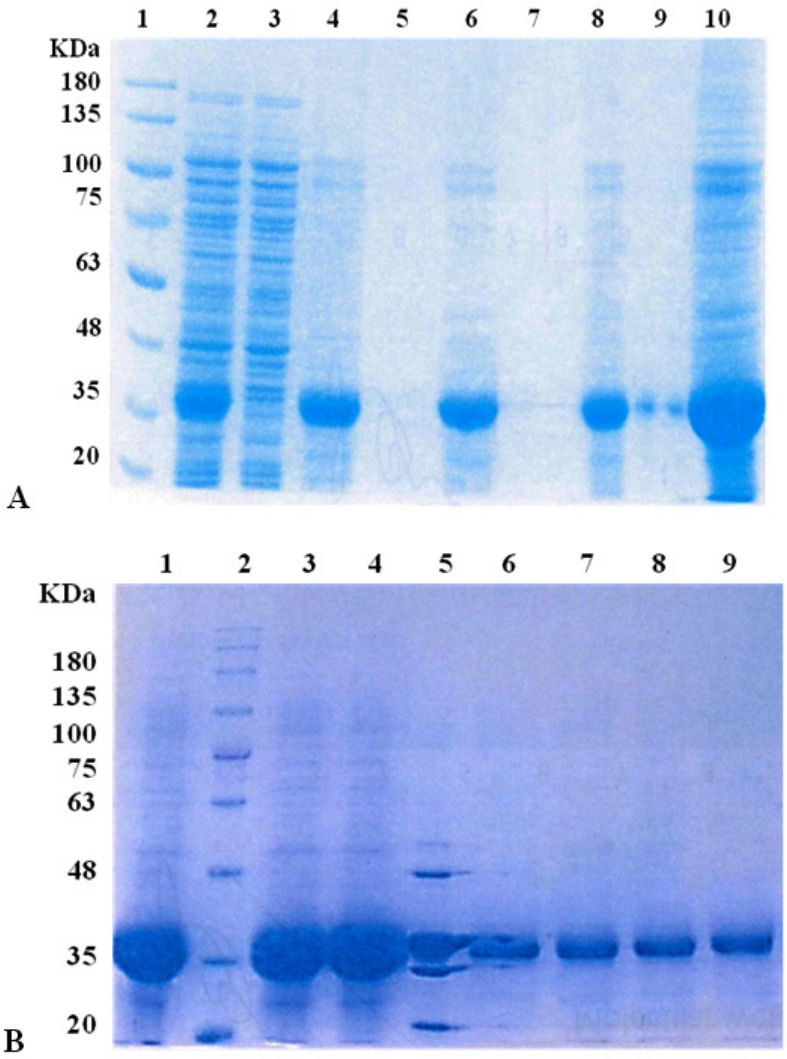
Table 5. Parameters monitored during fed-batch fermentation process

Time (h)	$O.D._{600}$	pH	Glucose	Time (h)	$O.D._{600}$	pH	Glucose
1	0.35	6.85	11.80	7	45.00	6.84	4.62
2	1.40	6.85	11.50	8	70.00	6.93	2.65
3	3.72	6.84	10.60	9	88.60	6.92	5.70
4	12.50	6.85	8.11	10	98.00	6.82	11.20
5	19.10	6.85	2.55	11	103.00	6.85	10.40
6	34.00	6.85	0.14	12	104.00	6.82	5.20

polyclonal antibodies against RBD and NSP9 using Western blotting technique to analyze the presence and correctness of the required epitopes. In Figure 7A and 7B, positive bands for NSP9 and RBD were observed when probed with in-house polyclonal antibody against NSP9 and the commercially available polyclonal antibody against RBD, respectively. However, both the blots also showed the presence of bands corresponding to dimeric and trimeric forms of the protein, which could be attributed to the polyclonal nature of the primary antibodies used in the assay.

DISCUSSION

COVID-19 is a global pandemic which has not been completely controlled. Newer cases of infection are being reported even today and mutant



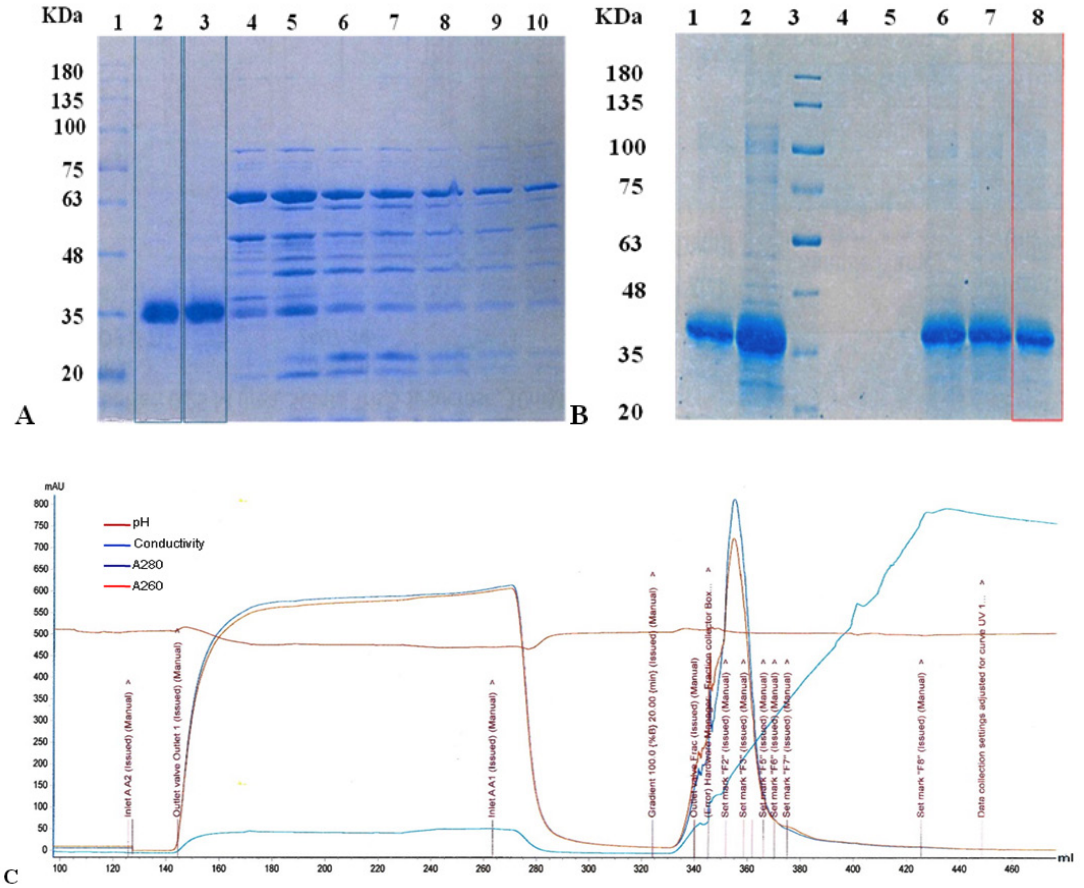
Lane 1: Molecular weight marker (20 kDa – 180 kDa); Lane 2: Homogenized Cell lysate; Lane 3: Supernatant; Lane 4: Cell pellet; Lane 5: Buffer wash 1 supernatant; Lane 6: Buffer wash 1 pellet; Lane 7: Buffer wash 2 supernatant; Lane 8: Buffer wash 2 pellet; Lane 9: WFI wash supernatant; Lane 10: WFI wash pellet. (B) SDS-PAGE analysis of solubilized inclusion bodies and refolded protein. Lane 1: Solubilized inclusion bodies; Lane 2: Molecular weight marker (20 kDa – 180 kDa); Lane 3: Solubilized supernatant; Lane 4: Solubilized supernatant (0.45µm filtered); Lane 5: Solubilized pellet; Lane 6: Refolding output at 0 h; Lane 7: Refolding output at 15 h; Lane 8: Refolding output at 15 h supernatant; Lane 9: Refolding output at 15 h, supernatant (0.45µm filter).

Fig. 4. (A) SDS-PAGE profile of purification of inclusion bodies

Table 6. Protein concentration by BCA Protein Assay and % step recovery for solubilized inclusion bodies, refolded protein, AEX chromatography run and TFF

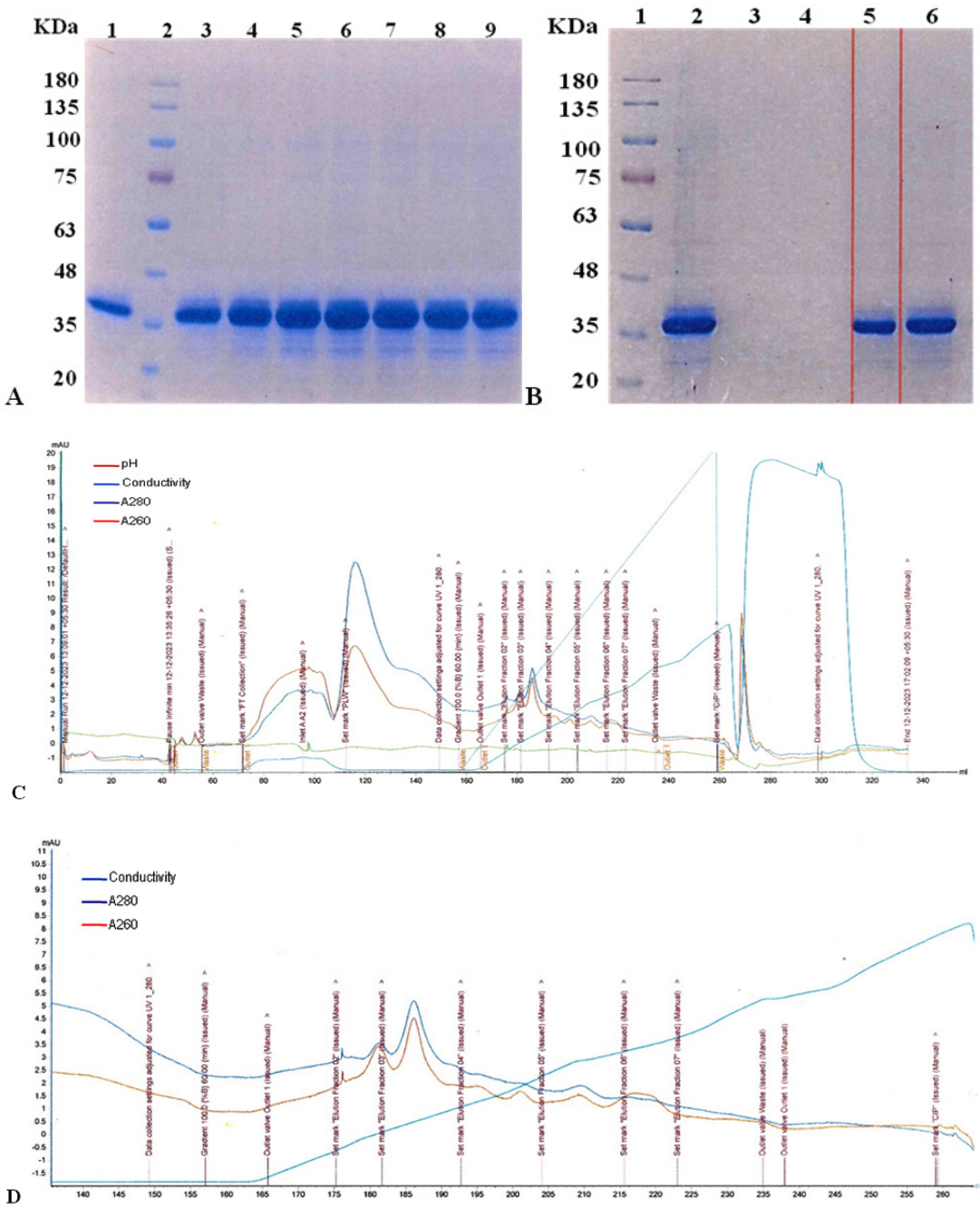
Sr. No.	Sample details	Volume (mL)	$O.D._{280}$	Protein Concentration (µg/mL)	Total Protein (µg)	% step recovery
1	Filtered Solubilized supernatant	50	0.985	498	24900	NA*
2	Refolding output after 15 h	500	0.102	45.2	22600	90.76%
3	AEX Flow-through	500	0.105	43.7	21850	96.68%
4	TFF Retentate	100	0.490	216.8	21680	99.22%

*NA = not applicable



Lane 1: Molecular weight marker (20 kDa – 180 kDa); Lane 2: AEX load; Lane 3: AEX flow-through; Lane 4: Elute fraction 1; Lane 5: Elute fraction 2; Lane 6: Elute fraction 3; Lane 7: Elute fraction 4; Lane 8: Elute fraction 5; Lane 9: Elute fraction 6; Lane 10: Elute fraction 7. (B) SDS-PAGE analysis of protein after TFF. Lane 1: AEX flow-through; Lane 2: Concentration retentate; Lane 3: Molecular weight marker (20 kDa – 180 kDa); Lane 4: Concentration permeate; Lane 5: TFF permeate; Lane 6: TFF retentate; Lane 7: TFF retentate after flushing; Lane 8: TFF retentate (0.45µm filtered). (C) Chromatogram of AEX chromatography. The chromatogram providing detailed information about A280, A260, pH and conductivity through the AEX chromatography run.

Fig. 5. (A) SDS-PAGE analysis of AEX chromatography run



Lane 1: CEX load; Lane 2: Molecular weight marker (20 kDa – 180 kDa); Lane 3: CEX Elute fraction 1; Lane 4: CEX Elute fraction 2; Lane 5: CEX Elute fraction 3; Lane 6: CEX Elute fraction 4; Lane 7: CEX Elute fraction 5; Lane 8: CEX Elute fraction 6; Lane 9: CEX Elute fraction 7. (B) SDS-PAGE analysis of protein after UFDF. Lane 1: Molecular weight marker (20 kDa – 180 kDa); Lane 2: Pooled CEX elute fractions (UFDF load); Lane 3: Ultrafiltration permeate; Lane 4: Diafiltration permeate; Lane 5: Final protein product (0.2 µm filtered); Lane 6: UFDF retentate after flushing. (C) Chromatogram of CEX chromatography. Detailed information about A280, A260, pH and conductivity through the CEX chromatography run was provided by the chromatogram. (D) Chromatogram of elution of CEX chromatography. Detailed information about A280, A260 and conductivity for the elution step of the CEX chromatography run was provided by the chromatogram.

Fig. 6. (A) SDS-PAGE profile of CEX chromatography run

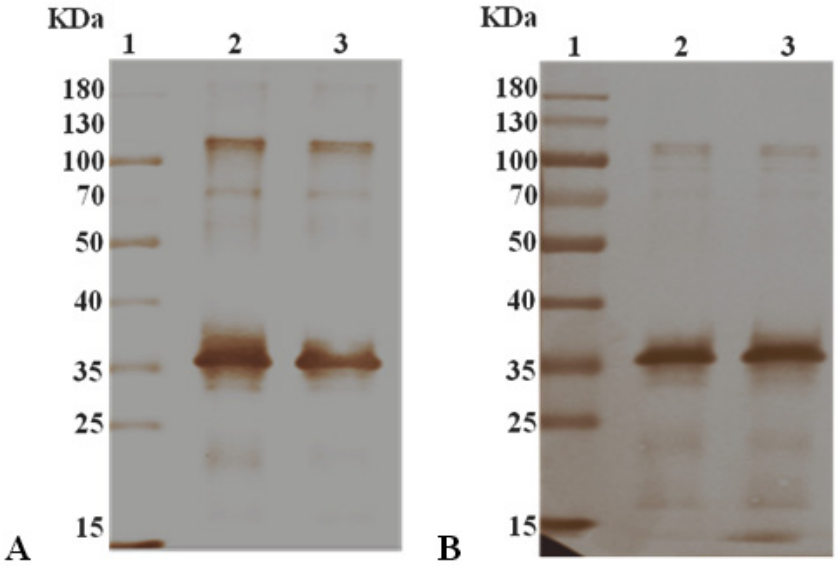
variants of the SARS-CoV2 virus are discovered almost every year. The global pandemic obligated fast track research and development of vaccine against the virus. Various research groups across the world were involved in developing a suitable vaccine for mass vaccination and increasing the immunity against the SARS-CoV2 virus[16,28]. The current study presented data for expression and purification of another such vaccine candidate, a multiple epitope protein-based vaccine using the most exploited structural protein, RBD[16,28], and a non-structural protein, NSP9 of SARS-CoV2.

E. coli is one of the most preferred hosts for expression of recombinant proteins due to its fast growth rate, ease of culture and purification of proteins. However, high levels of heterologous

Table 7. Protein concentration by BCA Protein Assay and % step recovery for CEX chromatography run

Sr. No.	Sample details	Volume (mL)	O.D.$_{280}$	Protein Concentration (µg/mL)	Total Protein (µg)	% step recovery
1	CEX Load	100	0.489	216.2	21620	NA*
2	CEX Flow-through	95	0.003	BQL**	NA*	NA*
3	CEX Post load wash	25	0.002	BQL**	NA*	NA*
4	CEX Elute Fraction 1	3	0.218	96.4	289.2	NA*
5	CEX Elute Fraction 2	3	0.320	141.5	424.4	NA*
6	CEX Elute Fraction 3	3	0.319	141.0	423.1	NA*
7	CEX Elute Fraction 4	3	0.340	150.3	451.0	NA*
8	CEX Elute Fraction 5	3	0.225	99.5	298.4	NA*
9	CEX Elute Fraction 6	3	0.226	99.9	299.8	NA*
10	CEX Elute Fraction 7	3	0.212	93.7	281.2	NA*
Total protein in pooled CEX Elute fractions					2467.1	11.41%

*NA= not applicable, **BQL=below quantification limit



Lane 1: Molecular weight marker (15 kDa – 180 kDa); Lane 2: 5 µg protein load; Lane 3: 3 µg protein load. (B) Western blot for purified protein with anti-RBD polyclonal antibody. Lane 1: Molecular weight marker (15 kDa – 180 kDa); Lane 2: 3 ìg protein load; Lane 3: 5 µg protein load.

Fig. 7. (A) Western blot for purified protein with anti-NSP9 polyclonal antibody

expression of recombinant proteins in *E. coli* can often result in formation of aggregates of the expressed protein, called inclusion bodies[29,30]. Although various strategies have been devised to obtain protein in the soluble form, inclusion bodies are advantageous since they contain >90% pure form of protein and hence allow for its easy purification[31,32].The present study showed that the expressed protein of ~35 kDa was achieved as inclusion bodies and the expression localization did not change when induction temperature was reduced. Localization of recombinant proteins in inclusion bodies even at low induction temperature is not an unseen phenomenon[25,27].

Various studies have been carried out to exploit inclusion bodies as a source of recombinant proteins. Agents like urea and guanidine hydrochloride are used for solubilization and refolding of protein from aggregates to achieve the bioactive form[26,27]. Researchers have shown that use of mild chaotropic agents like urea results in significantly large yields of active refolded protein[25,26,33]. The solubilization and refolding data of the current study demonstrated that use of urea increased the protein refolding efficiency, while minimizing protein loss (Table 6) corroborating with the data of earlier studies.

Protein purification was done using AEX and CEX to achieve the final purified protein. AEX chromatography is often used in the flow-through mode in order to remove host cell proteins, DNA and other impurities, especially for biotherapeutics like monoclonal antibodies[34–36]. Throughout the AEX chromatography, pH was maintained at 7.2, below the pI of the protein (8.32), thus working in the flow-through mode. The protein of interest was recovered in the flow-through and the extraneous cellular protein impurities that bound to the resin were removed as elutes. The overall purification process showed good recovery of protein at each step (Table 6). CEX chromatography is preferably used in the bind-and-elute mode[37]. A study by Adhikari *et al*. (2010) has shown that CEX chromatography can work as a single step purification method for basic proteins expressed in *E.coli*[38]. Similarly, owing to the basic pI of the protein (8.32) in this study, CEX chromatography worked as a bind-and-elute method for purification, although the recovery of protein was low (11%, Table 7). Since the purity of the protein was good, further analysis using western blotting with antibodies against NSP9 and RBD was performed. Western blots showed that the expressed protein had retained the epitopes of interest. Multiple bands were observed in the developed blots, which could be attributed to the polyclonal nature of the primary antibodies used. Another reason could be the conformational changes in the protein pertaining to the exposure to ion-exchange buffers. Ionic interactions have been reported to cause conformational changes in proteins, leading to formation of dimers, trimers or aggregates[37].

In summary, the present study provided data for the optimization of expression of the recombinant vaccine candidate and its purification. The data to prove the efficacy of the vaccine candidate is out of scope for this study and further studies involving *in vitro* and *in vivo* efficacy experiments will be pursued in the future.

CONCLUSION

While multiple vaccine candidates for COVID-19 have been developed, this study provided a novel candidate combining multiple epitopes in a single protein subunit vaccine. Multiple epitopes may lead to improved immune activation against COVID-19 infection. The current study presented data regarding the expression and purification of a recombinant multiple epitope vaccine using *E. coli* BL21 (DE3) as the expression host. The data further suggested that although the vaccine candidate was expressed as inclusion bodies, the process of refolding and purification of protein did not affect the correctness of epitopes. Western blotting analysis using polyclonal antibodies demonstrated that both the epitopes were retained in the final product.

ACKNOWLEDGEMENT

The authors acknowledge the generous gift of glycerol stock and Anti-SARS-CoV-2 NSP-9 polyclonal antibody from Aurigene Pharmaceutical Services Limited (APSL).

Conflict of Interest

The authors declare that they have no conflicts of interest regarding the publication of this research article.

Funding Source

This research did not receive any specific grant from funding agencies in the public, commercial or not- for-profit sectors.

Author's Contribution

Alok Singh contributed to the designing and execution of experiments, analysis of data and writing and compiling the manuscript. Prevesh Kumar, Diksha and Iqra Hasan carried out the proof-reading and editing of the manuscript. Navneet Verma approved the data and manuscript for publication.

Data Availability Statement

The manuscript incorporates all datasets produced or examined throughout this research study.

Ethics Approval Statement

Not Applicable.

REFERENCES

1. Pollard AJ, Bijker EM. A guide to vaccinology: from basic principles to new developments. *Nat Rev Immunol*. 2021;21(2):83-100. doi:10.1038/s41577-020-00479-7
2. Lahariya C. Vaccine epidemiology: A review. *J Family Med Prim Care*. 2016;5(1):7. doi:10.4103/2249-4863.184616
3. Tahamtan A, Charostad J, Hoseini Shokouh SJ, Barati M. An Overview of History, Evolution, and Manufacturing of Various Generations of Vaccines. *J Arch Mil Med*. 2017;In Press(In Press). doi:10.5812/jamm.12315
4. Baxter D. Active and passive immunity, vaccine types, excipients and licensing. *Occupational Medicine*. 2007;57(8):552-556. doi:10.1093/occmed/kqm110
5. Ahmad S, Baqar T, Kumar R. A Comprehensive Review on Types of Vaccines: From Classic to Cutting-Edge. *VVOA*. 2023;8(2):1-6. doi:10.23880/vvoa-16000164
6. Dai X, Xiong Y, Li N, Jian C. Vaccine Types. In: Kumar V, ed. *Vaccines - the History and Future*. IntechOpen; 2019. doi:10.5772/intechopen.84626
7. Al-Awwal N, Dweik F, Mahdi S, El-Dweik M, Anderson SH. A Review of SARS-CoV-2 Disease (COVID-19): Pandemic in Our Time. *Pathogens*. 2022;11(3):368. doi:10.3390/pathogens11030368
8. Lamers MM, Haagmans BL. SARS-CoV-2 pathogenesis. *Nat Rev Microbiol*. 2022;20(5):270-284. doi:10.1038/s41579-022-00713-0
9. Zhou H, Yang J, Zhou C, et al. A Review of SARS-CoV2: Compared With SARS-CoV and MERS-CoV. *Front Med*. 2021;8:628370. doi:10.3389/fmed.2021.628370
10. Low ZY, Zabidi NZ, Yip AJW, Puniyamurti A, Chow VTK, Lal SK. SARS-CoV-2 Non-Structural Proteins and Their Roles in Host Immune Evasion. *Viruses*. 2022;14(9):1991. doi:10.3390/v14091991
11. Wilamowski M, Hammel M, Leite W, et al. Transient and stabilized complexes of Nsp7, Nsp8, and Nsp12 in SARS-CoV-2 replication. *Biophysical Journal*. 2021;120(15):3152-3165. doi:10.1016/j.bpj.2021.06.006
12. Zhang C, Chen Y, Li L, et al. Structural basis for the multimerization of nonstructural protein nsp9 from SARS-CoV-2. *Mol Biomed*. 2020;1(1):5. doi:10.1186/s43556-020-00005-0
13. De O. Araújo J, Pinheiro S, Zamora WJ, Alves CN, Lameira J, Lima AH. Structural, energetic and lipophilic analysis of SARS-CoV-2 non-structural protein 9 (NSP9). *Sci Rep*. 2021;11(1):23003. doi:10.1038/s41598-021-02366-0
14. Zong S, Wu Y, Li W, et al. SARS-CoV-2 Nsp8 induces mitophagy by damaging mitochondria. *Virologica Sinica*. 2023;38(4):520-530. doi:10.1016/j.virs.2023.05.003
15. Littler DR, Liu M, McAuley JL, et al. A natural product compound inhibits coronaviral replication in vitro by binding to the conserved Nsp9 SARS-CoV-2 protein. *Journal of Biological Chemistry*. 2021;297(6):101362. doi:10.1016/j.jbc.2021.101362
16. Rahman MdM, Masum MdHU, Wajed S, Talukder A. A comprehensive review on COVID-19 vaccines: development, effectiveness, adverse effects, distribution and challenges. *VirusDis*. 2022;33(1):1-22. doi:10.1007/s13337-022-00755-1
17. Vikkurthi R, Ansari A, Pai AR, et al. Inactivated whole-virion vaccine BBV152/Covaxin elicits robust cellular immune memory to SARS-CoV-2 and variants of concern. *Nat Microbiol*. 2022;7(7):974-985. doi:10.1038/s41564-022-01161-5
18. Choi TJ, Geletu TT. High level expression and purification of recombinant flounder growth hormone in E. coli. *Journal of Genetic Engineering and Biotechnology*. 2018;16(2):347-355. doi:10.1016/j.jgeb.2018.03.006
19. Laemmli UK. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. Published online 1970.
20. Coomassie Brilliant Blue staining solution. *Cold Spring Harbor Protocols*. 2007;2007(2):pdb.rec10727. doi:10.1101/pdb.rec10727

21. Coomassie destaining solution. *Cold Spring Harbor Protocols*. 2007;2007(4):pdb.rec10932. doi:10.1101/pdb.rec10932
22. 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. *Commun Biol*. 2020;3(1):214. doi:10.1038/s42003-020-0939-8
23. Eggenreich B, Wurm DJ, Rajamanickam V, Klausser R, Slouka C, Spadiut O. High pressure homogenization is a key unit operation in inclusion body processing. *Journal of Biotechnology*. 2020;324:100022. doi:10.1016/j.btecx.2020.100022
24. Singh A, Upadhyay V, Panda AK. Solubilization and Refolding of Inclusion Body Proteins. In: García-Fruitós E, ed. *Insoluble Proteins*. Vol 1258. Methods in Molecular Biology. Springer New York; 2015:283-291. doi:10.1007/978-1-4939-2205-5_15
25. Singh A, Upadhyay V, Singh A, Panda AK. Structure-Function Relationship of Inclusion Bodies of a Multimeric Protein. *Front Microbiol*. 2020;11:876. doi:10.3389/fmicb.2020.00876
26. Upadhyay AK, Singh A, Mukherjee KJ, Panda AK. Refolding and purification of recombinant L-asparaginase from inclusion bodies of E. coli into active tetrameric protein. *Front Microbiol*. 2014;5. doi:10.3389/fmicb.2014.00486
27. Zhao T, Huang H, Tan P, et al. Enhancement of Solubility, Purification, and Inclusion Body Refolding of Active Human Mitochondrial Aldehyde Dehydrogenase 2. *ACS Omega*. 2021;6(18):12004-12013. doi:10.1021/acsomega.1c00577
28. Sabitha S, Shobana N, Prakash P, et al. A Review of Different Vaccines and Strategies to Combat COVID-19. *Vaccines*. 2022;10(5):737. doi:10.3390/vaccines10050737
29. Rosano GL, Morales ES, Ceccarelli EA. New tools for recombinant protein production in *Escherichia coli*/ : A 5-year update. *Protein Science*. 2019;28(8):1412-1422. doi:10.1002/pro.3668
30. Tripathi NK. Production and Purification of Recombinant Proteins from *Escherichia coli*. *ChemBioEng Reviews*. 2016;3(3):116-133. doi:10.1002/cben.201600002
31. Bhatwa A, Wang W, Hassan YI, Abraham N, Li XZ, Zhou T. Challenges Associated With the Formation of Recombinant Protein Inclusion Bodies in Escherichia coli and Strategies to Address Them for Industrial Applications. *Front Bioeng Biotechnol*. 2021;9:630551. doi:10.3389/fbioe.2021.630551
32. De Marco A, Ferrer-Miralles N, Garcia-Fruitós E, et al. Bacterial inclusion bodies are industrially exploitable amyloids. *FEMS Microbiology Reviews*. 2019;43(1):53-72. doi:10.1093/femsre/fuy038
33. Singh A, Upadhyay V, Upadhyay AK, Singh SM, Panda AK. Protein recovery from inclusion bodies of Escherichia coli using mild solubilization process. *Microb Cell Fact*. 2015;14(1):41. doi:10.1186/s12934-015-0222-8
34. Rathore AS, Hebbi V. Ion Exchange Chromatographic Methods for Purification of Therapeutic Antibodies. In: Houen G, ed. *Therapeutic Antibodies*. Vol 2313. Methods in Molecular Biology. Springer US; 2022:179-186. doi:10.1007/978-1-0716-1450-1_10
35. Silva RJS, Moleirinho MG, Moreira AS, et al. A Flow-Through Chromatographic Strategy for Hepatitis C Virus-Like Particles Purification. *Processes*. 2020;8(1):85. doi:10.3390/pr8010085
36. Masuda Y, Ogino Y, Yamaichi K, Takahashi Y, Nonaka K, Wakamatsu K. The prevention of an anomalous chromatographic behavior and the resulting successful removal of viruses from monoclonal antibody with an asymmetric charge distribution by using a membrane adsorber in highly efficient, anion exchange chromatography in flow through mode. *Biotechnology Progress*. 2020;36(3):e2955. doi:10.1002/btpr.2955
37. Guo J, Xu X. Chapter 25 - Conformational changes of biomolecules in ion-exchange chromatography. In: Nesterenko PN, Poole CF, Sun Y, eds. *Ion-Exchange Chromatography and Related Techniques*. Elsevier; 2024:521-534. doi:10.1016/B978-0-443-15369-3.00020-1
38. Adhikari S, Manthena PV, Sajwan K, Kota KK, Roy R. A unified method for purification of basic proteins. *Analytical Biochemistry*. 2010;400(2):203-206. doi:10.1016/j.ab.2010.01.011