Cloning, Improved Expression and Purification of Invasion Plasmid Antigen D (IpaD): An Effector Protein of Enteroinvasive *Escherichia Coli* (EIEC)

Sudeshna Halder
National Institute of Technology Durgapur
Namita Jaiswal
National Institute of Technology Durgapur
Nibedita Mahata (nibedita.mahata@bt.nitdgp.ac.in)
National Institute of Technology Durgapur https://orcid.org/0000-0002-6026-1187

Research Article

Keywords: Cloning, Enteroinvasive Escherichia coli, Protein stability, Purification

Posted Date: January 18th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1224024/v1

Abstract

The widespread increase in broad-spectrum antimicrobial resistance is making it more difficult to treat gastrointestinal infections. Enteroinvasive *Escherichia coli* is a prominent etiological agent of bacillary dysentery, invading via the fecal-oral route and exerting virulence on the host via the type III secretion system. IpaD, a surface-exposed protein on the T3SS tip that is conserved among EIEC and *Shigellae*, may serve as a broad antigen for bacillary dysentery protection. For the first time, we present an effective framework for improving the expression level and yield of IpaD in the soluble fraction for easy recovery, as well as ideal storage conditions, which may aid in the development of new protein therapies for gastrointestinal infections in the future. To achieve this, uncharacterized full length *IpaD* gene from EIEC was cloned into pHis-TEV vector and induction parameters were optimised for enhanced expression in the soluble fraction. After affinity-chromatography based purification, 61% pure protein with a yield of 0.33 mg per litre of culture was obtained. The purified IpaD was kept the secondary structure with predominant α-helical structure at 4°C, -20°C, and -80°C using 5% sucrose as cryoprotectants during storage which is the fundamental parameter for protein-based therapeutics.

Introduction

The Global Enteric Multicenter Study (GEMS) reports bacillary dysentery as one of the leading causes of morbidity and mortality in young children under 5 years of age in developing nations such as India, Bangaladesh, Sri Lanka, Nepal, Bhutan, and Myanmar (Troeger et al. 2018). According to the Global Burden of Disease 2016 study, it is ranked as the eighth leading cause of mortality, responsible for more than 1·6 million deaths worldwide (Khalil et al. 2018; Troeger et al. 2018). Enteroinvasive *Escherichia coli* (EIEC), a pathotype of *E. coli*, is the leading cause of inflammation and ulceration of the intestinal epithelium in humans, followed by bloody and mucoid diarrhoea. The pathogenicity mechanism for both EIEC and *Shigella* is similar, hence the physiological and biochemical differentiation of the infection caused by them is quite challenging (Lan et al. 2004; Van den Beld & Reubsaet 2012; Van den Beld et al. 2019). Thus, the evaluation of the actual burden of EIEC induced bacillary dysentery is challenging. The EIEC infections are generally sporadic, however few well studied cases of outbreak such as the 1970s outbreak in United States (Marier et al. 1973), the 2012 outbreak in Italy (Escher et al. 2014) and two linked outbreaks of United Kingdom in 2014 (Newitt et al. 2016) reiterate that causative agents of bacillary dysentery is not limited to the species of *Shigella* genus, but EIEC significantly partakes in such events.

Diarrhoea has been a long-standing priority for the World Health Organization, yet no vaccine for bacillary dysentery currently exists (Hosangadi et al. 2019). Thus, it is one of the leading causes of antibiotic prescription and consumption among children in low- and middle-income countries, leading to resistance against the major third-generation antibiotics (Eltai et al. 2020; Rogawski McQuade et al. 2020). Further, the limited understanding of EIEC and antimicrobial therapy necessitates the development of effective vaccines against diarrhoeal pathogens, which may contribute to the ancillary benefits such as reducing antibiotic exposure and resistance (Nguyen et al. 2005). The mechanism of virulence in many Gram-negative pathogens is facilitated by the type III secretion system (T3SS). It serves as a conduit for transfer of bacterial effector proteins into the host cell facilitating the host invasion and infection by the bacteria. IpaD, a conserved 37 kDa hydrophilic protein present on T3SS needle tip controls translocator and effector protein secretion in EIEC is the first bacterial protein that

interacts with the host cell (Espina et al. 2006). The immunogenic nature of IpaD confirmed by immune profiling, together with its conserved nature across the species of the genus *Shigella* and EIEC makes it promising vaccine candidate (Martinez-Becerra et al. 2012; Ndungo et al. 2018; Turbyfill et al. 1998).

Numerous studies have focused on understanding the immunoprotective behaviour of IpaD as vaccine candidate. Jahantigh *et. al*. reported that IpaD shows a highly protective humoral response in guineapig as IgA titter level significantly increased upon nasal administered of chitosan nanofibrous membrane containing N-terminal region of (Jahantigh et al. 2014). Another study with N-terminal IpaD loaded trimethylated chitosan nanoparticles induced increase in IgG and IgA levels in guinea pigs and exhibited protective behaviour against *Shigella* infection (Akbari et al. 2019). These studies clearly enphasize the importance of IpaD protein based vaccine conjugated to different carriers. It is also to note that in order to design and develop protein based formulations, it is critical to optimize the parameters for enhanced expression and recombinant purification of IpaD in *E. coli*. But to date, there is no report on optimization of parameters for expression, purification and storage to ensure structural and functional stability of protein, which are crucial for the development of any biopharmaceutical formulation.

In the present study, we report the molecular cloning of *IpaD* into pHis-TEV vector, optimization of expression parameters such as inducer concentration and temperature to improve the yield of IpaD protein. Here, we also report effective purification strategies of IpaD by using simple Ni-NTA affinity chromatography. To the best of our knowledge, this study is the first attempt to clone *IpaD* gene of EIEC and subsequent successful expression and purification of the protein in *E. coli* BL21 (DE3) strain was achieved. At the optimized IPTG concentration of 0.5 mM, resulted in ~37% IpaD expression with ~67% protein distribution in the soluble fraction-a relatively high quantity. After Ni-NTA affinity chromatography purification 61% pure protein was obtained. We further studied the optimal temperature and buffer for storage so that structural and functional integrity of the protein remains unaffected for its potential application as biopharmaceuticals. Our results demonstrate that the purified IpaD can be stored at 4°C, -20°Cand -80°C with no structural loss, confirmed by CD analysis.

Materials And Methods

Materials

All chemicals were used in this study was analytical grade and used as received without any further purification. All solutions were prepared in milli-Q ultrapure water of resistivity not less than 18.2 MΩ cm^{-1}.

Bacterial strains and plasmid

The bacterial strain Enteroinvasive *Escherichia coli* was gifted from NICED, Kolkata. All plasmid and expression vector used in this study were listed in Table no. 1A and 1B. All strains were handled in Biosafety level-II facility (Thermo Fisher Scientific A2 1300 Series).

Extraction of DNA and PCR amplification of *IpaD* gene

On Luria Bertani Broth Miller (Himedia, Mumbai, India), a single isolated colony of Enteroinvasive *E.coli* was cultured for 12 hours at 37°C with vigorous shaking at 200 rpm (Hazen et al. 2016). To collect the cell pellet, the culture was centrifuged at 13000 rpm for 5 minutes. The alkaline lysis procedure was used to isolate the

plasmid DNA (Feliciello & Chinali 1993).Primers were generated using the IDT oligo analyzer tool and synthesized by IDT Technologies (India) for the amplification of the *IpaD* gene. To insert amplified *IpaD* gene into the pHis-TEV plasmid Vector, forward and reverse primers (Table No. 2) were constructed containing EcoRI and XhoI restriction endonuclease enzyme sites respectively. Gradient polymerase chain reaction was performed using the thermal Cycler (Biorad T100) to optimise the annealing temperature. A gradient temperature of 57.5°C to 61.5°C was set. The reaction was initiated by heating the reaction mixture at 95°C for 3 minutes, followed by 30 cycles of denaturation at 95°C for 30 s, annealing for 30 s, and elongation at 68°C for 1 minute, and finally extension at 72°C for 10 minutes. Using the DNA agarose gel Electrophoresis Apparatus (Tarsons), the PCR product was analyzed in a 0.8% agarose (Sigma Aldrich) gel run at 70 V for 1 hour in 1X TAE (Tris-acetate-EDTA) buffer and bands were visualized using the UV Transilluminator (Himedia). The PCR product was purified using a Qiagen PCR Purification kit according to the manufacturer's instructions. The Nano-Drop Spectrophotometer (Eppendorf) was used to determine the DNA concentration.

Molecular cloning into expression plasmid and verification of the insert

The purified PCR product, i.e. *IpaD* gene and pHis-TEV plasmid vector were digested with EcoRI and XhoI restriction enzyme in digestion buffer (New England Biolabs) at 37°C for 2 hours as directed by the manufacturer and analysed on a 0.8% agarose gel. The T4 DNA ligase enzyme (New England Biolabs) was used to ligate the digested gene and plasmid vector, which was done at 16°C overnight as per the manufacturer's protocol. The resulted ligated product (pHis-TEV-IpaD) was transformed into *Escherichia coli*-DH5α competent cell following standard $CaCl_2$ heat shock transformation protocol (Li et al. 2010) and the full construct of pHis-TEV containing IpaD gene was shown in (Fig. 1A). The positive colony was confirmed by double digestion analysis using EcoRI and XhoI restriction enzyme and colony PCR which were analyzed in a 0.8% agarose gel. The separated gene and vector fragmented part were visualized using the Gel Doc (Biorad) system and analyzed by Image Lab (Biorad) Software. The colony was maintained Luria Bertani agar (Himedia) plate containing ampicillin (100µg/ml) (Himedia) as selectable marker. For DNA sequencing, plasmid DNA was isolated from the positive colony using Qiagen Plasmid isolation kit as per manufacturer protocol. The sequencing was done by Integrated DNA Technologies (India) by the Sanger Sequencing method. NCBI Blast was used to analyse the sequence results.

IpaD protein expression optimization

Optimization of inducer concentration

Escherichia coli-BL21 (DE3) cells containing pHis-TEV-IpaD plasmid was grown overnight in Luria Bertani Broth supplemented with ampicillin (100 µg/ml). Five flasks containing fresh LB media were inoculated with overnight grown culture (1:100) and incubated at 37°C, 150 rpm, until the OD_{600} reached between ~0.7-0.8. The cultures were induced individually by five different isopropylthio-β-galactoside (IPTG) Concentrations, respectively 0.05mM, 0.25mM, 0.5mM, 1 mM and 2 mM for 18 hours at 15□C with 150 rpm. After induction, cells were harvested by centrifugation at 8000 rpm for 10 min at 4°C then the resultant pellet was resuspended in lysis buffer (20 mM Tris (Himedia) (pH 8.0), 500 mM NaCl (Himedia),10 mM imidazole (Himedia) and 5% sucrose (Himedia) in a 1:100 lysis buffer:culture volume ratio. Then cells were disrupted by sonication (5 cycles, 15-second pulse with 1-minute interval). To prevent protein degradation, 1mM Phenyl methyl sulfonyl fluoride (PMSF, Sigma Aldrich) and 1mg/ml lysozyme (Himedia) were added before sonication. The soluble and

insoluble part were separated by centrifugation. The soluble fraction was collected in the supernatant portion by centrifuged at 20,000 rpm for 30 min at 4°C, and the insoluble part containing inclusion bodies was collected as pellet at bottom of the tube. The pellet was washed two times with a lysis buffer to remove any contaminant of soluble part. 8 M urea was added to the pellet to solubilize the insoluble inclusion bodies and boiled for 15 min, harvested by centrifugation. Protein expression and solubility were analysed by SDS -PAGE using ImageJ Software following two formulas.

$$\text{Expression level} = \frac{S}{I}$$

$$\text{Solubility Level} = \frac{S''}{(S'' + P)}$$

Whereas, S is the amount of IpaD protein and I is the total protein after induced by IPTG; S" is the amount of the IpaD in supernatant fraction and P is the amount of total protein in pellet fraction

Optimization of post induction temperature

Escherichia coli-BL21 (DE3) cells harbouring pHis-TEV-IpaD plasmid were grown overnight with antibiotic containing LB media. From the overnight culture, three flasks containing fresh media were inoculated and kept at 37 ℃ with 150 rpm until the OD_{600} reached ~0.7-0.8. After that, the effect of temperature on the expression of IpaD, culture was induced with different IPTG concentration and kept at three different temperatures such as 10℃, 15℃ and 37℃ for 18 hours. After the completion of incubation time, cells were harvested and disrupted by sonication, and expression was analyzed by densitometry analysis.

Purification of IpaD protein

Escherichia coli BL21 (DE3) containing pHis-TEV-IpaD plasmid was grown for 18 hours at 15° C with 150 rpm after induction with 0.5 mM IPTG. Then induced cells were collected by centrifugation and disrupted by sonication. Consequential supernatants were separated by centrifugation and filtered through a 0.45µm syringe filter unit (Himedia) to the trace amount of debris. Then, IpaD protein was purified by Ni-NTA affinity chromatography using the 5 ml His trap column (GE Health Care) as per manufacturer instructions with minor modification. Briefly, the column was equilibrated with 10 column volume (CV) equilibration buffer (20 mM Tris (pH 8.0), 500 mM NaCl, 10 mM imidazole) containing 5% Sucrose. The supernatant was passed through the column four times for proper binding, and flow-through was collected. Afterwards, the column was washed twice with 10ml CV wash buffer (20 mM Tris (pH 8.0), 500 mM NaCl, 5% Sucrose) with two different concentration of imidazole (20 mM imidazole, 60 mM imidazole) and wash fraction was also collected. The protein was eluted by applying the 5 ml elution buffer (20 mM Tris (pH 8.0), 500 mM NaCl, 5% Sucrose) with three different concentration of imidazole (150 mM imidazole, 250 mM imidazole, and 500 mM imidazole). Then eluted fractions were immediately diluted with 1:1 ratio dilution buffer (20 mM Tris (pH 8.0), 500 mM NaCl) and were dialyzed to remove out the salt and imidazole against 1X TBS Buffer (20 mM Tris (pH 7.4), 150 mM NaCl, 5% sucrose) for 6 hours. The his tagged-IpaD protein was incubated with TEV protease overnight at 4°C (Sigma Aldrich) to remove the His-tag as described in the user manual provided by manufacturer. Further, the IpaD

protein was applied to the His-trap column, and the flow-through fraction was collected as purified protein. Protein concentration was estimated using a Bradford reagent (Biorad) and analyzed by SDS -PAGE.

Analysis of protein stability

Purified IpaD protein (1mg/ml) was stored in cryotubes containing 1XTBS Buffer (20mM Tris, 150 mM NaCl pH 7.4) with 5% sucrose at 25°C (as room temperature), 4°C, -20°C and -80°C for 15 days. The secondary structure of IpaD protein was then analysed by Far UV CD.

Analytical procedure

SDS PAGE analysis and densitometry analysis

For SDS–PAGE analysis, 1X loading buffer (1M Tris-HCl (pH 6.8), 0.8 gm SDS, 10% glycerol, 14.7M β-mercaptoethanol,0.5M EDTA, 8mg bromophenol Blue) was added with the protein samples and was heated for 5 min at 95 C. Protein samples were run in 15% resolving and 4% stacking gels for 2 hours at 120 V (Biorad mini protein system). The gel was stained with Coomassie blue (50% methanol, 10% acetic acid and 0.1% Coomassie brilliant blue R-250) thereafter destanied with (45% methanol and 10% acetic acid). A prestained ladder (Biorad) was used as standard. For densitometry analysis, the area under the curve (AUC) of the protein was analyzed from the recorded image by using molecular imager (Biorad), and images were analyzed by ImageJ software.

Western blot analysis

For western blotting after resolving the IpaD protein sample was transferred onto the PVDF membrane (Millipore) using semidry transfer apparatus (Biorad) for 1 hour with constant voltage at 20V. After complete transfer membrane was washed with TBST buffer (20 mM Tris pH7.5, 150mM NaCl and 0.1% (v/v) Tween-20 (Sigma) and then blocked with blocking buffer (3% BSA in TBST buffer) for 1 hour at room temperature with gentle shaking. The membrane was then incubated with mouse anti-IpaD monoclonal antibody (1:1000, Abbexa, USA) in TBST containing 1% BSA overnight at 4 °C. Next, the membrane was washed for three times with TBST then incubated with anti-mouse secondary antibody conjugated with HRP (IgG-HRP) (1:5000; Invitrogen, USA) for 1 hour at room temperature in 3% (w/v) BSA in TBST. After washing thrice with TBST, immunoblot was developed in the presence of a chemiluminescence substrate (Bio-Rad).

Size exclusion chromatography

IpaD was further concentrated using Amicon ultrafilter (Millipore) and purified by size exclusion chromatography using a Superdex 200 10/300 GL column (GE Healthcare). Column was equilibrated using the mobile phase as TBS Buffer (20mM Tris, 150mM NaCl pH-7.4) containing 5% sucrose. The flow rate was maintained at 0.4ml/min, and UV detection was set at 280 nm. 500 µg protein was loaded with injection volume of 500µl. Eluted purified was analysed by SDS PAGE.

Circular dichroism analysis

For secondary structure analysis, Far-UV CD was performed using J-815 Spectrophotometer (Jasco). Three Spectra was measured between the 250-200 region with scan speed 50nm/min and bandwidth 1 nm. 1X TBS

(20mM Tris, 150mM NaCl pH-7.4) buffer containing 5% sucrose was used as control for the analysis.

Statistical analysis

Statistical analysis was performed by using Student's t test to calculate significance of differences. All experiments were accomplished at least three replicates and the data are presented as the mean ± standard deviation (SD).The values of *p < 0.05 represents significant, **p < 0.01and ***p < 0.001 represents very significant.

Results

Molecular cloning of *IpaD*

Cloning of full length *IpaD* gene from *EIEC* was carried out by PCR amplification using the specific primers (Table 2). For amplification, different annealing temperatures were set based on the melting temperature of primers, and the optimal annealing temperature was determined to be 59.5°C as evident from the agarose gel band intensity which was maximum for lane 3 (Fig. 1B). To rule out false positive amplification of the polymerase chain reaction, a negative control (without template) was run (Fig.1B, Lane 5).The inclusion of insert (*IpaD* gene, 999 bp) was validated by double digestion with EcoRI and XhoI before sequencing. Agarose gel electrophoresis of double digested products of pHis-TEV-IpaD showed two bands located near ~1kbp and ~6 kbp, depicting gene and vector respectively (Fig.1C, Lane 6). The undigested empty vector showed three distinct bands depicting nicked, linear, and supercoiled structures respectively in agarose gel (Fig.1C, Lane 2); whereas, a single band depicting linear structure was present for the double digested empty vector (Fig.1C, Lane 3). Furthermore, the insertion of *IpaD* gene was again confirmed by conventional colony PCR. Three colonies were picked randomly as templates and colony PCR was performed. Agarose gel results showed a band at ~ 1kbp for one colony (Fig.S1, Lane 2), which was also confirmed to be positive by double digestion experiments (Fig. 1C). A positive control was run to validate PCR reaction using the amplified product as template (Fig.S1, Lane 5). After getting the nucleotide sequence the blast result shows that our sequence was 99% identical with *E. coli* strains and 98-96% identical with *Shigella* subtypes, but with 100% coverage for both types (Fig.S2). Hence, it might be concluded that the insertion of *IpaD* gene into the pHis-TEV vector was successful.

Optimization of inducer concentration for IpaD protein expression

As the expression of IpaD protein was under the control of the T7 promoter, IPTG was used to induce protein expression in *E. coli* BL21 (DE3). The pHis-TEV-IpaD plasmid was transformed into *E. coli* BL21 (DE3) strain for protein expression. The primary parameter to be optimized for expression of proteins in any bacterial expression system is the inducer concentration, IPTG. In this study, we tested five different IPTG concentrations, 0.05mM, 0.25mM, 0.5mM, 1mM, and 2 mM with three different post-induction temperatures 10°C, 15°C, and 37°C respectively for a fixed post-induction period of 18 hours. Post induction, the IpaD content in the total cell extract, soluble and insoluble fractions at 10°C (Fig 2A), 15°C (Fig. 2B) and 37°C (Fig. 2C), with varying concentration of IPTG were analysed by SDS – PAGE densitometry. An uninduced *E.coli* BL21 (DE3) containing pHis-TEV-IpaD plasmid served as the negative control for each experiment set. Representative SDS-PAGE was showing expression level of IpaD protein induced with 0.05 mM (Fig. 2A, 2B, 2C; Lane: 2-4), 0.25mM (Fig. 2A, 2B, 2C; Lane: 5-7), 0.5mM (Fig. 2A, 2B, 2C; Lane: 9-11), 1 mM (Fig. 2A, 2B, 2C; Lane: 12-14), and 2 mM

(Fig. 2A, 2B, 2C; Lane: 16 -18) IPTG concentration. Expression level of IpaD protein in total cell extract induced with different IPTG concentration was shown in the Fig. 2A, 2B, 2C (Lane: 2, 5, 9, 12, and 16). Expression level of IpaD protein in the soluble fraction induced with different IPTG concentration was shown in the Fig. 2A, 2B, 2C (Lane: 3, 6, 10, 13, and 17). Expression level of IpaD protein in the insoluble fraction induced with different IPTG concentration was shown in the Fig. 2A, 2B, 2C (Lane: 4, 7, 11, 14, and 18). Effect of temperature and IPTG concentration on the expression and solubility level of expressed IpaD protein were quantified by densitometry analysis (Fig. 2D and 2E) using ImageJ software.

The expression level of IpaD with 0.5 mM IPTG concentration was 37% and increased up to 52% for 1.0 mM IPTG concentration (Fig. 2D). However, the protein solubility substantially decreased from 67% at 0.5 mM of IPTG to 40% at 1.0 mM of IPTG concertation (Fig. 2E). According to the results, as depicted in Fig 2D and 2E, induction by other IPTG concentrations less amount of IpaD was expressed and solubility was decreased. So, among the five IPTG concentration variation 0.5 mM IPTG concentration was chosen for further study because this concentration displayed maximum soluble expression.

Optimization of post induction temperature for IpaD protein expression

Further, to prevent protein aggregation and achieve maximal soluble expression of IpaD proteins in *E. coli* three different post induction temperatures was varied .To this end, three flasks containing LB medium were cultivated under optimized conditions. As shown in Fig 2D and 2E, at induction temperature of 37°C and 10°C, both expression and solubility decreased but significantly solubility was increased at 15°C.Therefore, the optimised post-induction temperature was preferred 15°C.

Purification of IpaD protein

Nickel column affinity chromatography was used for purification of the IpaD protein containing 6×His in its N-terminal. The steps involved in purification of IpaD is depicted by Fig 3A. A prominent protein band of ~40 kDa corresponding to IpaD present in soluble fraction was observed on SDS-PAGE (Fig. 3B, Lane 2). For separation and recombinant purification IpaD, the supernatant was allowed to bind by passing it through Ni-NTA column. Removal of impurities and non-specific protein was achieved by extensive washing with 20 mM first and then with 60 mM imidazole which breaks non-specific binding of protein. (Fig. S3, Lane 4 and Lane 5). In this study, the imidazole concentration in the elution buffer was varied as follows: 150 mM, 250mM and 500mM. It is observed that imidazole concentrations of 250 mM and 500 mM in elution buffer resulted complete elution of bound IpaD (Fig S4, Lane 8, Lane 7). However, at 150 mM imidazole concentration the amount of IpaD eluted was far less (Fig S4, Lane 9). Together, these results suggested that 250 mM imidazole can be used in elution buffer for successful purification of IpaD. IpaD was efficiently purified as the eluate consisted mainly of a single prominent protein band of ~40 kDa (Fig. 3B, Lane 5). In each purification step, percentage of purification and yield were calculated by densitometry analysis of SDS-PAGE using Image J (Table 3) (Burgess, 2009).

His tag was removed by TEV protease cleavage as per manufacturer's instruction. SDS PAGE analysis of tag-cleaved IpaD does not exhibit any major difference with His tag protein probably due to the small size of the His tag ~3kDa (Fig 3C, Lane 3 and 4). The densitometry analysis showed that the band corresponded to 60.84% with 2.82 mg pure protein from 1000ml culture. Furthermore, the expressed protein was purified by size-exclusion chromatography. The chromatogram showed that the protein was eluted at retention volume 23.70 ml (Fig 4A). The eluted protein was analysed by SDS PAGE (Fig 4B, Lane1). Quantitative analysis was revealed

that each litre of culture yields 0.33 mg pure IpaD protein. Finally, western blot analysis with mouse monoclonal anti-IpaD antibody confirmed the purified protein as IpaD (Fig 4C, Lane 1).

Purified IpaD stability and integrity analysis

Considering all these factors, the purified IpaD protein was subjected to dialysis against TBS pH 7.4 and then stored at four different temperatures (4°C, 25°C, -20°C and -80°C) for 15 days in TBS Buffer (20 mM Tris, 150 mM NaCl, pH 7.4) with 5% sucrose as stabilizing agent to determine the ideal storage condition. After 15 days of incubation at different temperatures, the secondary structure analysis by CD spectrum revealed that the samples stored in 4°C, -20°C and -80°C were structurally stable, detailed discussion follows in the subsequent section.

Circular Dichroism Analysis and Secondary Structure determination

To the best of our knowledge, neither the crystal structure of full-length EIEC IpaD protein nor its secondary structure content has been reported to date. Therefore, we studied the secondary structure content of IpaD using far UV Circular Dichroism (CD) spectroscopy. To estimate the secondary structural contents of purified IpaD, far UV CD spectroscopic analysis was performed. CD spectrum of IpaD showed a negative peak at 222 nm, and 208 nm confirming the presence of the α- helix; presence of antiparallel β- sheet were confirmed by presence of the negative peak at 218 nm (Fig. 5A) (Greenfield 2006). From the BestSel, the secondary structural components of IpaD consisted of 44% α- helix, 25.1% β- sheet, and 11.3% turn for samples stored at 4°C; 35% α- helix, 16% β- sheet and 10.2% turn at -20°C; 26.3 % α- helix, 21% β- sheet and 10% turn at -80°C. However at room temperature *i.e.* 25°C, percentage of α- helix content IpaD protein drastically decreased to 1.9%, β- sheet increased to 36.4% and 15.3% of turns were observed as shown in Fig.5B. The drastic reduction of α- helixcal structures along with high variation in β- sheet and turns clearly indicate that protein was not stable at room temperature. However, the percentage of secondary structural components viz., α- helix, β- sheet and turns were found to be in close approximation for 4°C and - 20°C. For -80°C stored IpaD samples, α- helix percentage slightly decreased in comparison to 4°C and - 20°C (Fig 5B). CD analysis revealed that IpaD secondary structure predominantly consisted of α- helix along with a good percentage of β sheet and turn.

Discussion

IpaD is one of most imperative Shigella virulence agents and is essential for pathogenesis .IpaD has a crucial role in TTSS, as it controls the secretion of IpaB and IpaC (Jahantigh et al. 2014). Moreover, IpaD have ability to excites both mucosal and humoral immunity (Akbari et al. 2019). The surface exposed nature and high degree of conservation among the group of *EIEC* and *Shigella* species IpaD makes itself a promising antigens for bacillery dysentery (Martinez-Becerra et al. 2012). Since the successful over-expression of IpaD could be very valuable, the stability of the purified IpaD on different storage temperatures has great importance. In this study, the aim was to optimise the soluble expression of recombinant IpaD in the *E. coli* host and achieve a simple purification of the soluble form. Two fundamental parameters were tested: the inducer concentration and the induction temperature. No such extensive study on the recombinant expression of IpaD in an *E.coli* expression system and the storage temperature effect on purified IpaD has been reported previously. Therefore to best of our knowledge, this is the first attempt to develop a broader picture of IpaD production, stability and integrity of

the protein on storage, it is crucial to enhance our understanding of the soluble expression, purification of IpaD in *E. coli* and storage condition.

To maximize expression of the *IpaD* gene, modified pET21d (pHis-TEV) vector with an innate 6X His-tag at the N-terminal and a TEV protease site allowing TEV protease enzyme directed cleavage of His tag from the recombinant IpaD protein, was used. The PCR amplified *IpaD* gene product was incorporated next to the T7 promoter and lac operator (Fig. 1A), which aids in the overexpression of the IpaD protein (Liu & Naismith 2009) as well as limiting uninduced expression (Shilling et al. 2020). N-terminal His tag allowed easy protein purification, without affecting the protein's structural or functional integrity and inhibits protein aggregation in the column by limiting intermolecular interactions (Booth et al. 2018; Seetaraman Amritha et al. 2020). Final confirmation of successful clones was done by nucleotide sequencing using the sanger sequencing method with universal primers (Mardis 2017). Subsequently, the sequence obtained was compared against the available complete genome sequence of EIEC in NCBI Blast site for cloning confirmation (Johnson et al. 2008).

To optimize the inducer concentration for overexpression of *IpaD* gene to achieve a high yield of protein in the soluble form. To accomplish this objective, *E.coli* BL21(DE3), containing T7 RNA polymerase under the control of lacZ promoter inducible with IPTG (Heyde & Nørholm 2021) was chosen as host system due to its several advantages like rapid growth (Francis & Page 2010) with high cell densities (Rosano & Ceccarelli 2014), can be maintained in low-cost media (Huang et al. 2012), is capable of producing varied therapeutic proteins and finally enhances the yield of target protein (Angius et al. 2018). Addition of IPTG was done when the OD_{600} value reached to 0.7 -0.8 depicting bacterial exponential growth phase, containing of highest number of cells to maximize the yield of soluble protein (Larentis et al. 2014).

For IPTG concentration of 0.5mM, the level of IpaD expression was 37% and increased up to 52% for 1.0 mM IPTG concentration (Fig. 2D). However, the protein solubility substantially decreased from 67% at 0.5 mM of IPTG to 40% at 1.0 mM of IPTG concertation (Fig. 2E) at 15°C. With increase in the IPTG concentration beyond 0.5 mM, IpaD protein were predominantly localized into the inclusion body *i.e.* level of solubility was decreased at higher concentration of IPTG (Fig. 2E). Additionally, higher IPTG concentration had undesirable effects on cell growth and soluble protein yield which may be attributed to the increase of metabolic load because target protein overexpression by bacterial cells and results in interrupted metabolic activity and overall decrease of protein yield (Paramanik & Thakur 2011). This observation is supported by previous reports recommending the use of low IPTG concentration for induction owing to the following reasons i) higher concentrations are toxic to cell, ii) expensive, iii) potential problem of product insolubility, iv) It induces growth inhibition and cell lysis (Donovan et al. 1996; Volontè et al. 2008).

Post induction temperature also plays a decisive role in the solubility of the target protein. Lower induction temperature improves the target protein expression level and also enhances its content in the soluble fraction. Lower temperatures also reduces the cell stress (Jhamb & Sahoo 2012), inhibits protein aggregation and helps in proper protein folding (Nguyen et al. 2019; Vera et al. 2006). As shown in Fig. 2D and 2E, at induction temperature of 37°C and 10°C, both expression and solubility decreased. It is well established that at higher induction temperature, hydrophobic interaction reactions enhanced, thus preventing the S-S bond formation which eventually leads to the formation of aggregated unfolded protein with unstable tertiary structure in inclusion body (de Groot & Ventura 2006; Schein 1989). At induction temperature of 10°C, the expression and solubility was also affected as bacterial growth was barred due to loss of membrane fluidity and enzyme

activity (Song et al. 2012). Together these results established that the optimum induction temperature was 15°C for the expression of IpaD protein at 0.5 mM of IPTG concertation, among the three temperatures studied. Earlier, it is reported that the induction temperature 15°C-30°C was the suboptimum temperature range for protein overexpression as the solubility and expression level substantially increased as this temperature range supports production of correctly folded polypeptide and decrease the heat denaturation and heat shock proteases (Song et al. 2012). Therefore, together these results suggested that a low IPTG concentration of 0.5 mM and the induction temperature of 15°C were the optimum conditions required for high yield production of appropriately folded protein in soluble fraction.

In the present study, Ni-NTA column was used for the purification of expressed IpaD protein based on the interaction between the immobilized divalent metal ion (Ni^{+2}) and imidazole group of histidine (Bornhorst & Falke 2000). The pI of IpaD from Expasy server was calculated to be 5.32 (Wilkins et al. 1999). Therefore, pH of the wash buffer and elution buffer pH was adjusted to 8.0 (> pI) to have a net negatively charged protein (Novák & Havlícek 2016) which aids protein retention on the column (Lee et al. 2008; Ueda et al. 2003) for better purification. A wash buffer with low concentration of imidazole and a high concentration of salt (500 mM) was used to reduce non-specific binding caused by disruption of the hydrophobic interaction between the protein and the resin (Bornhorst & Falke 2000; Kielkopf et al. 2020). Imidazole concentration is also an important factor for protein elution because of its competitive nature towards the metal against histidine, which helps to elute all proteins without difficulty (Lee et al. 2008). Hence, the optimization of imidazole concentration in the elution buffer is a critical parameter for any protein purification study. Hence, the optimization of imidazole concentration in the elution buffer is a critical parameter for any protein purification study. In the present study we have seen that at 250 mM and 500 mM imidazole concentration almost the same amount of protein eluted but previous study reported that imidazole have a negative impact on protein stability (Walter, 2020). So, from this point of view we have chosen 250 mM imidazole concentration in elution buffer. Tag-less IpaD was further purified using Ni-NTA column and was eluted with 60 mM imidazole (Nguyen et al. 2019). In the present study we have calculated band intensity and getting a total 2.82 mg with ~61% pure protein after TEV cleavage and the purity profile was characterized by size exclusion chromatography, confirmed by western blot analysis with anti-IpaD monoclonal antibody. Therefore, from this study we concluded that by following this simple purification method IpaD protein was purified successfully.

It is essential to determine an optimal storage condition for purified proteins as the structural integrity of proteins may be easily compromised due to aggregation during purification, shipping and storage process (Jain et al. 2021; Rathore & Rajan 2008). Loss of structural integrity subsequently leads to loss of therapeutic action of purified protein (Chen et al., 2017). Protein is typically stored at refrigerated temperatures ranging from 2°C to 8°C, at -20°C, and at extreme freezing temperatures starting from -80°C depending on the type and stability of protein (Simpson 2010). Furthermore, apart from storage temperature the right stabilizer is another crucial factor to maintain protein integrity. Recent reports show that 5% sucrose used as a stabilizing agent in both liquid and freezing states help stabilize protein structure by maintaining the hydration layer (Flood et al. 2016; Kendrick et al. 1997; Olsson et al. 2020; Pelliccia et al. 2016). Buffers are also a crucial elements of protein storage conditions for maintaining protein stability and functionality (Kamerzell et al. 2011; Ugwu & Apte 2004). For instance, the amine-based structure of Tris-HCl buffer is capable of preserving constant pH in both aqueous and frozen states (Kim et al. 2021; Taha & Lee 2010) whereas, the pH of phosphate buffer is drastically lowered under freezing conditions (Kim et al. 2021). The results presented in this work confirmed that the purified IpaD

protein maintains its stability and integrity for in TBS buffer with 5% sucrose as cryoprotectant, which is promising for the use of IpaD as a protein formulation. CD is an extensively used analytical technique for secondary structure determination of purified proteins (Greenfield 2006). CD analysis of proteins convey secondary conformational information such as α-helix, β-sheet, turn, and random coil (Greenfield 2006). Our result also consistent with the characteristics of CD spectrum that α-helix show negative peaks at 222 and 208 nm and a positive peak at 193 nm. Similarly, antiparallel β-pleated sheets (β-sheets) show characteristic negative peak at 218 nm and a positive peak at 195 nm, while that for disordered random coils have a positive peak above 210 nm and a negative peak near 195 nm (Greenfield 2006). BestSel webserver was used to calculate percentage of the different secondary structures of the protein from CD spectroscopy data (Micsonai et al. 2018). It is a reliable method for the accurate scrutiny of α- helix and β- sheet because other algorithms predict the cumulative α- helix structure (Micsonai et al. 2018; Micsonai et al. 2015). The stability at wide temperatures viz., 4°C, -20°C and -80°C corroborates with the fact that most of protein therapeutics that are manufactured as liquid formulation are mainly stored at 2°C to 8°C (regular cold chain) or in a deepfreeze state i.e -20°C to -80°C implying that as a protein therapeutics, IpaD could efficiently maintain its structural and functional stability (Yu et al. 2021).

Conclusion

In conclusion, this study is the first attempt to clone *IpaD* gene of EIEC and subsequent successful expression and purification of the IpaD protein in *Escherichia coli* BL21 (DE3) strain. The EIEC *IpaD* gene was successfully cloned into the pHis-TEV vector by optimizing annealing temperature. We have also optimised the expression parameters, such as IPTG concentration and induction temperature – the key players for any recombinant production in soluble form and purification of protein efficiently using simple techniques. Our results validate improved purification as evident from high- yield and high-purity protein quality. Another important fundamental parameter that we addressed herein is the storage temperature, prerequisite of any protein biologics for further applications. Structural analysis of purified IpaD protein showed that it efficiently retained its secondary structure when stored at 4°C, -20°C and -80°C, implying that as a protein therapeutics, IpaD could efficiently maintain its structural and functional stability. Therefore these findings represent a significant step towards the soluble production in *E. coli*, provide optimal purification method and storage condition of IpaD, a potential therapeutic candidate for bacillary dysentery.

Declarations

Acknowledgments:

Authors would like to express their heartfelt gratitude to the Director of NIT Durgapur for his constant support and encouragement. Authors are also thankful to Dr. Nripen Chanda, Senior Scientist, Material Processing and Microsystem Laboratory, CSIR–Central Mechanical Engineering Research Institute, Durgapur-713209, India for kind help.

Funding

This material is based upon work supported by National Institute of Technology, Durgapur, India.

Compliance with ethical standards

No human and animals are used or harmed in this study.

Conflict of interest

Authors have no conflict of interests.

References

1. Akbari MR, Saadati M, Honari H et al (2019) IpaD-loaded N-trimethyl chitosan nanoparticles can efficiently protect guinea pigs against Shigella flexneri. Iranian Journal of Immunology 16:212–224. https://doi.org/10.22034/IJI.2019.80272
2. Angius F, Ilioaia O, Amrani A et al (2018) A novel regulation mechanism of the T7 RNA polymerase based expression system improves overproduction and folding of membrane proteins. Sci Rep 8:1–11. https://doi.org/10.1038/s41598-018-26668-y
3. Booth WT, Schlachter CR, Pote S et al (2018) Impact of an N-terminal polyhistidine tag on protein thermal stability. ACS omega 3:760–768. https://doi.org/10.1021/acsomega.7b01598
4. Bornhorst JA, Falke JJ (2000) Purification of proteins using polyhistidine affinity tags. Methods Enzymol 326:245–254. https://doi.org/pmc/articles/PMC2909483
5. Burgess RR (2009) Preparing a purification summary table. Methods Enzymol 463:29–34. https://doi.org/10.1016/S0076-6879(09)63004-4
6. Chen YC, Smith T, Hicks RH, Doekhie A, Koumanov F, Wells SA, Edler KJ, Van Den Elsen J, Holman GD, Marchbank KJ, Sartbaeva A (2017) Thermal stability, storage and release of proteins with tailored fit in silica. Sci Rep 7:1–8. https://doi.org/10.1038/srep46568
7. de Groot NS, Ventura S (2006) Effect of temperature on protein quality in bacterial inclusion bodies. FEBS Lett 580:6471–6476. https://doi.org/10.1016/J.FEBSLET.2006.10.071
8. Donovan RS, Robinson CW, Glick B (1996) Optimizing inducer and culture conditions for expression of foreign proteins under the control of the lac promoter. J Ind Microbiol 16:145–154. https://doi.org/10.1007/BF01569997
9. Eltai NO, Al Thani AA, Al Hadidi SH et al (2020) Antibiotic resistance and virulence patterns of pathogenic Escherichia coli strains associated with acute gastroenteritis among children in Qatar. BMC Microbiol 20:1–12. https://doi.org/10.1186/s12866-020-01732-8
10. Escher M, Scavia G, Morabito S et al (2014) A severe foodborne outbreak of diarrhoea linked to a canteen in Italy caused by enteroinvasive Escherichia coli, an uncommon agent. Epidemiology & Infection 142:2559–2566. https://doi.org/10.1017/S0950268814000181
11. Espina M, Olive AJ, Kenjale R et al (2006) IpaD localizes to the tip of the type III secretion system needle of Shigella flexneri. Infect Immun 74:4391–4400. https://doi.org/10.1128/IAI.00440-06
12. Feliciello I, Chinali G (1993) A modified alkaline lysis method for the preparation of highly purified plasmid DNA from Escherichia coli. Anal Biochem 212:394–401. https://doi.org/10.1006/abio.1993.1346
13. Flood A, Estrada M, McAdams D et al (2016) Development of a freeze-dried, heat-stable influenza subunit vaccine formulation. PLoS ONE 11:e0164692. https://doi.org/10.1371/journal.pone.0164692

14. Francis DM, Page R (2010) Strategies to optimize protein expression in E. coli. Current protocols in protein science 61:5.24. 1-5.24. 29. https://doi.org/10.1002/0471140864.PS0524S61
15. Greenfield NJ (2006) Using circular dichroism spectra to estimate protein secondary structure. Nat Protoc 1:2876–2890. https://doi.org/10.1038/nprot.2006.202
16. Hazen TH, Leonard SR, Lampel KA et al (2016) Investigating the relatedness of enteroinvasive Escherichia coli to other E. coli and Shigella isolates by using comparative genomics. Infect Immun 84:2362–2371. https://doi.org/10.1128/IAI.00350-16
17. Heyde SA, Nørholm MH (2021) Tailoring the evolution of BL21 (DE3) uncovers a key role for RNA stability in gene expression toxicity. Communications Biology 4:1–9. https://doi.org/10.1038/s42003-021-02493-4
18. Hosangadi D, Smith PG, Kaslow DC et al (2019) WHO consultation on ETEC and Shigella burden of disease, Geneva, 6–7th April 2017: Meeting report. Vaccine 37:7381–7390. https://doi.org/10.1016/j.vaccine.2017.10.011
19. Huang C-J, Lin H, Yang X (2012) Industrial production of recombinant therapeutics in Escherichia coli and its recent advancements. J Ind Microbiol Biotechnol 39:383–399. https://doi.org/10.1007/S10295-011-1082-9
20. Jahantigh D, Saadati M, Ramandi MF et al (2014) Novel intranasal vaccine delivery system by chitosan nanofibrous membrane containing N-terminal region of IpaD antigen as a nasal Shigellosis vaccine, Studies in Guinea pigs. J Drug Deliv Sci Technol 24:33–39. https://doi.org/10.1016/S1773-2247(14)50005-6
21. Jain K, Salamat-Miller N, Taylor K (2021) Freeze–thaw characterization process to minimize aggregation and enable drug product manufacturing of protein based therapeutics. Sci Rep 11:1–9. https://doi.org/10.1038/s41598-021-90772-9
22. Jhamb K, Sahoo DK (2012) Production of soluble recombinant proteins in Escherichia coli: effects of process conditions and chaperone co-expression on cell growth and production of xylanase. Bioresour Technol 123:135–143. https://doi.org/10.1016/J.BIORTECH.2012.07.011
23. Johnson M, Zaretskaya I, Raytselis Y et al (2008) NCBI BLAST: A better web interface. Nucleic Acids Res 1:5–9. https://doi.org/10.1093/nar/gkn201
24. Kamerzell TJ, Esfandiary R, Joshi SB et al (2011) Protein–excipient interactions: Mechanisms and biophysical characterization applied to protein formulation development. Adv Drug Deliv Rev 63:1118–1159. https://doi.org/10.1016/j.addr.2011.07.006
25. Kendrick BS, Chang BS, Arakawa T et al (1997) Preferential exclusion of sucrose from recombinant interleukin-1 receptor antagonist: role in restricted conformational mobility and compaction of native state. Proceedings of the National Academy of Sciences 94:11917-11922. https://doi.org/10.1073/pnas.94.22.11917
26. Khalil IA, Troeger C, Blacker BF et al (2018) Morbidity and mortality due to shigella and enterotoxigenic Escherichia coli diarrhoea: the Global Burden of Disease Study 1990–2016. Lancet Infect Dis 18:1229–1240. https://doi.org/10.1016/S1473-3099(18)30475-4
27. Kielkopf CL, Bauer W, Urbatsch IL (2020) Purification of polyhistidine-tagged proteins by immobilized metal affinity chromatography. Cold Spring Harbor Protocols 2020:pdb prot 102194. https://doi.org/10.1101/PDB.PROT102194

28. Kim A-Y, Kim H, Park SY et al (2021) Development of a Potent Stabilizer for Long-Term Storage of Foot-and-Mouth Disease Vaccine Antigens. Vaccines 9:252. https://doi.org/10.3390/vaccines9030252
29. Lan R, Alles MC, Donohoe K et al (2004) Molecular evolutionary relationships of enteroinvasive Escherichia coli and Shigella spp. Infect Immun 72:5080–5088. https://doi.org/10.1128/IAI.72.9.5080-5088.2004
30. Larentis AL, Nicolau JFMQ, dos Santos Esteves G et al (2014) Evaluation of pre-induction temperature, cell growth at induction and IPTG concentration on the expression of a leptospiral protein in E. coli using shaking flasks and microbioreactor. BMC Res Notes 7:1–13. https://doi.org/10.1186/1756-0500-7-671
31. Lee JJ, Bruley DF, Kang KA (2008) Effect of pH and imidazole on protein C purification from Cohn fraction IV-1 by IMAC. in: *Oxygen Transport to Tissue XXVIII*, (Eds.) M. DJ, B. DF, H. D.K, Springer. Boston, MA, pp. 61-66. https://doi.org/10.1007/978-0-387-71764-7_9
32. Li X, Sui X, Zhang Y et al (2010) An improved calcium chloride method preparation and transformation of competent cells. Afr J Biotechnol 9:8549–8554. https://doi.org/10.4314/ajb.v9i50
33. Liu H, Naismith JH (2009) A simple and efficient expression and purification system using two newly constructed vectors. Protein Exp Purif 63:102–111. https://doi.org/10.1016/j.pep.2008.09.008
34. Mardis ER (2017) DNA sequencing technologies: 2006–2016. Nat Protoc 12:213–218. https://doi.org/10.1038/nprot.2016.182
35. Marier R, Wells J, Swanson R et al (1973) An outbreak of enteropathogenic Escherichia coli foodborne disease traced to imported French cheese. The Lancet 302:1376–1378. https://doi.org/10.1016/S0140-6736(73)93335-7
36. Martinez-Becerra FJ, Kissmann JM, Diaz-McNair J et al (2012) Broadly protective Shigella vaccine based on type III secretion apparatus proteins. Infect Immun 80:1222–1231. https://doi.org/10.1128/IAI.06174-11
37. Micsonai A, Wien F, Bulyáki É et al (2018) BeStSel: a web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra. Nucleic Acids Res 46:W315–W322. https://doi.org/10.1093/nar/gky497
38. Micsonai A, Wien F, Kernya L et al (2015) Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. Proceedings of the National Academy of Sciences 112:E3095-E3103. https://doi.org/10.1073/pnas.1500851112
39. Ndungo E, Randall A, Hazen TH et al (2018) A novel Shigella proteome microarray discriminates targets of human antibody reactivity following oral vaccination and experimental challenge. Msphere 3:e00260–e00218. https://doi.org/10.1128/MSPHERE.00260-18
40. Newitt S, MacGregor V, Robbins V et al (2016) Two linked enteroinvasive Escherichia coli outbreaks, Nottingham, UK, June 2014. Emerg Infect Dis 22:1178. https://doi.org/10.3201/eid2207.152080
41. Nguyen MT, Prima MJ, Song J-A et al (2019) Prokaryotic soluble overexpression and purification of oncostatin M using a fusion approach and genetically engineered E. coli strains. Sci Rep 9:1–13. https://doi.org/10.1038/s41598-019-50110-6
42. Nguyen TV, Le Van P, Le Huy C et al (2005) Detection and characterization of diarrheagenic Escherichia coli from young children in Hanoi. Vietnam Journal of clinical microbiology 43:755–760. https://doi.org/10.1128/JCM.43.2.755-760.2005
43. Novák P, Havlícek V (2016) Protein extraction and precipitation, Proteomic Profiling and Analytical Chemistry, Elsevier. https://doi.org/10.1016/B978-0-444-63688-1.00004-5

44. Olsson C, Zangana R, Swenson J (2020) Stabilization of proteins embedded in sugars and water as studied by dielectric spectroscopy. Physical Chemistry Chemical Physics 22:21197–21207. https://doi.org/10.1039/d0cp03281f
45. Paramanik V, Thakur M (2011) Overexpression of mouse estrogen receptor-β decreases but its transactivation and ligand binding domains increase the growth characteristics of E. coli. Mol Biotechnol 47:26–33. https://doi.org/10.1007/S12033-010-9308-Z
46. Pelliccia M, Andreozzi P, Paulose J et al (2016) Additives for vaccine storage to improve thermal stability of adenoviruses from hours to months. Nat Commun 7:1–7. https://doi.org/10.1038/ncomms13520
47. Rathore N, Rajan RS (2008) Current perspectives on stability of protein drug products during formulation, fill and finish operations. Biotechnol Prog 24:504–514. https://doi.org/10.1021/BP070462H
48. Rogawski McQuade ET, Shaheen F, Kabir F et al (2020) Epidemiology of Shigella infections and diarrhea in the first two years of life using culture-independent diagnostics in 8 low-resource settings. PLoS Negl Trop Dis 14:e0008536. https://doi.org/10.1371/JOURNAL.PNTD.0008536
49. Rosano GL, Ceccarelli EA (2014) Recombinant protein expression in Escherichia coli: advances and challenges. Front Microbiol 5:172. https://doi.org/10.3389/FMICB.2014.00172
50. Schein CH (1989) Production of soluble recombinant proteins in bacteria. Bio/Technology 7:1141–1149. https://doi.org/10.1038/nbt1189-1141
51. Seetaraman Amritha TM, Mahajan S, Subramaniam K et al (2020) Cloning, expression and purification of recombinant dermatopontin in Escherichia coli. PLoS ONE 15:e0242798. https://doi.org/10.1371/journal.pone.0242798
52. Shilling PJ, Mirzadeh K, Cumming AJ et al (2020) Improved designs for pET expression plasmids increase protein production yield in Escherichia coli. Communications biology 3:1–8. https://doi.org/10.1038/s42003-020-0939-8
53. Simpson RJ (2010) Stabilization of proteins for storage. Cold Spring Harbor Protocols 2010:pdb. top79. https://doi.org/10.1101/pdb.top79
54. Song JM, An YJ, Kang MH et al (2012) Cultivation at 6–10 C is an effective strategy to overcome the insolubility of recombinant proteins in Escherichia coli. Protein Exp Purif 82:297–301. https://doi.org/10.1016/J.PEP.2012.01.020
55. Taha M, Lee M-J (2010) Interactions of TRIS [tris (hydroxymethyl) aminomethane] and related buffers with peptide backbone: thermodynamic characterization. Physical Chemistry Chemical Physics 12:12840–12850. https://doi.org/10.1039/c0cp00253d
56. Troeger C, Blacker BF, Khalil IA et al (2018) Estimates of the global, regional, and national morbidity, mortality, and aetiologies of diarrhoea in 195 countries: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Infect Dis 18:1211–1228. https://doi.org/10.1016/S1473-3099(18)30362-1
57. Turbyfill KR, Mertz JA, Mallett CP et al (1998) Identification of epitope and surface-exposed domains of Shigella flexneri invasion plasmid antigen D (IpaD). Infection and immunity 66:1999-2006. https://doi.org/10.1128/iai.66.5.1999-2006.1998
58. Ueda E, Gout P, Morganti L (2003) Current and prospective applications of metal ion–protein binding. J Chromatogr A 988:1–23. https://doi.org/10.1016/S0021-9673(02)02057-5

59. Ugwu SO, Apte SP (2004) The effect of buffers on protein conformational stability. Pharm Technol 28:86–109. https://doi.org/10.1014/j.addr.2014.07.009
60. Van den Beld M, Reubsaet F (2012) Differentiation between Shigella, enteroinvasive Escherichia coli (EIEC) and noninvasive Escherichia coli. European journal of clinical microbiology & infectious diseases 31:899–904. https://doi.org/10.1007/s10096-011-1395-7
61. Van den Beld MJ, Warmelink E, Friedrich AW et al (2019) Incidence, clinical implications and impact on public health of infections with Shigella spp. and entero-invasive Escherichia coli (EIEC): Results of a multicenter cross-sectional study in the Netherlands during 2016–2017. BMC Infect Dis 19:1–12. https://doi.org/10.1186/S12879-019-4659-Y
62. Vera A, González-Montalbán N, Garcia-Fruitós E et al (2006) Low growth temperatures improve the conformational quality of aggregation prone recombinant proteins in both soluble and insoluble E. coli cell fractions. Microb Cell Fact 5:1–2. https://doi.org/10.1186/1475-2859-5-S1-P7
63. Volontè F, Marinelli F, Gastaldo L et al (2008) Optimization of glutaryl-7-aminocephalosporanic acid acylase expression in E. coli. Protein Exp Purif 61:131–137. https://doi.org/10.1016/J.PEP.2008.05.010
64. Walter BM, Szulc A, Glinkowska MK (2020) Reliable method for high quality His-tagged and untagged E. coli phosphoribosyl phosphate synthase (Prs) purification. Protein Exp Purif 169:1–7. https://doi.org/10.1016/j.pep.2020.105587
65. Waugh DS (2011) An overview of enzymatic reagents for the removal of affinity tags. Protein Exp Purif 80:283–293. https://doi.org/10.1016/J.PEP.2011.08.005
66. Wilkins M, Gasteiger E, Bairoch A et al (1999) Protein identification and analysis tools in the ExPASy server. https://doi.org/10.1385/1-59259-584-7:531. Methods Mol Biol531-552
67. Yu YB, Briggs KT, Taraban MB et al (2021) Grand challenges in pharmaceutical research series: ridding the cold chain for biologics. Pharm Res 38:3–7. https://doi.org/10.1007/s11095-021-03008-w

Tables

Table 1 (A) *Escherichia coli* strains and (B) plasmid used in this study

1A: *Escherichia coli* strains

SL NO.	*E.coli* STRAIN	DESCRIPTION	SOURCE
1.	Enteroinvasive *E.coli*	Mother Strain	NICED, Kolkata,(India)
2.	*Escherichia coli* -DH5α	Plamid Construction	MTCC, Chandigarh(India)
3	*Escherichia coli*-BL21(DE3)	Protein Expression	MTCC, Chandigarh(India)

1B: Plasmid

SL NO.	PLASMID	DESCRIPTION	SOURCE
1.	pHis-TEV	Amp^R, T7 promoter with 6X-His-TEV protease site	Biobharti, Kolkata,(India)

Table 2 Primers used to clone *IpaD* gene from EIEC

Primer Name	Primer Sequence	Restriction enzyme Name	enzyme Sequence
Forward	5'GATCGGAATTCATGAATATAACAACTCTGACTAATAGTATTTCCACC3'	EcoRI	GAATTC
Reverse	5'GATCGCTCGAGTCAGAAATGGAGAAAAAGTTTATCTGTATCTG3'	**XhoI**	CTCGAG

Table 3 Purification of IpaD protein expressed in *Escherichia coli* BL21 (DE3).

Total protein quantified by Bradford reagent method using BSA as a standard protein. a. Purity of the protein of interest was determined by densitometry using ImageJ. b. The yield at each step was defined as the amount of IpaD at that step divided by the amount in the first step, considered 100%.

Purification Step	Total Protein (mg)	IpaD (mg)	% Purity([a])	% Yield([b])
Supernatant(Soluble fraction)	203	15.43	20.80	100
1st IMAC	35.1	6.21	49.22	26.84
2nd IMAC (after TEV cleavage)	7.0	2.82	60.82	12.19
SEC purified	0.36	0.33	100	1.41

Figures

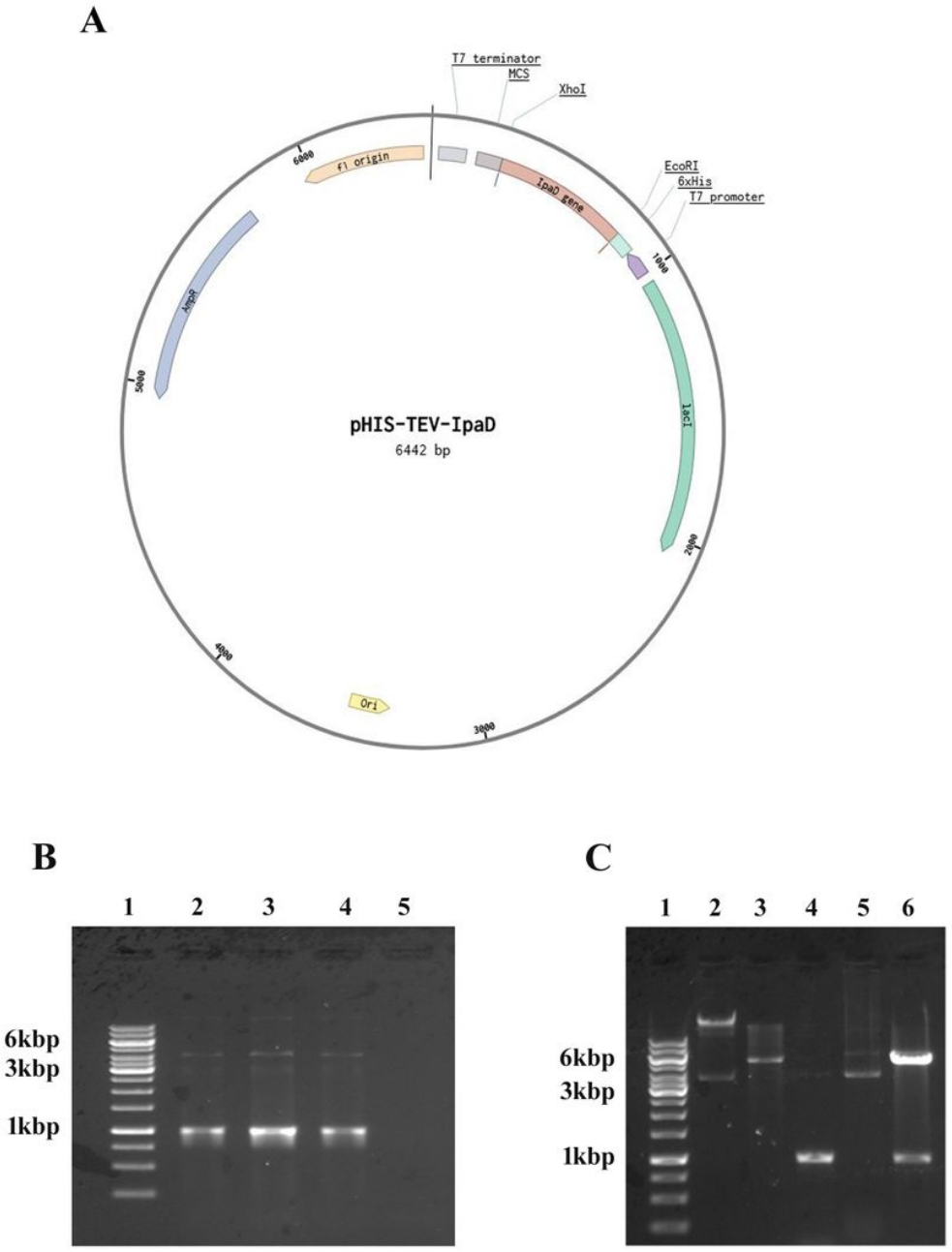

Figure 1

PCR amplification and cloning confirmation of *IpaD* gene

A) Agarose gel electrophoresis of gradient PCR of *IpaD* gene at three different annealing temperatures. Lane 1: 1 Kb DNA Ladder; Lane 2: Amplified product at 57.5°C; Lane 3: Amplified product at 59.5°C; Lane 4: Amplified product at 61.5°C; Lane 5: Negative Control (PCR master mixture without gene product).

B) Restriction digestion of recombinant Vector (pHis-TEV-IpaD). Lane 1: 1Kb DNA Ladder; Lane 2: Empty pHis-TEV vector without digestion; Lane 3: Empty pHis-TEV vector digested with EcoRI and XhoI; Lane 4: *IpaD* gene

product; Lane 5: Transformed product (pHis-TEV-IpaD); Lane 6: Transformed product digested with EcoRI and XhoI

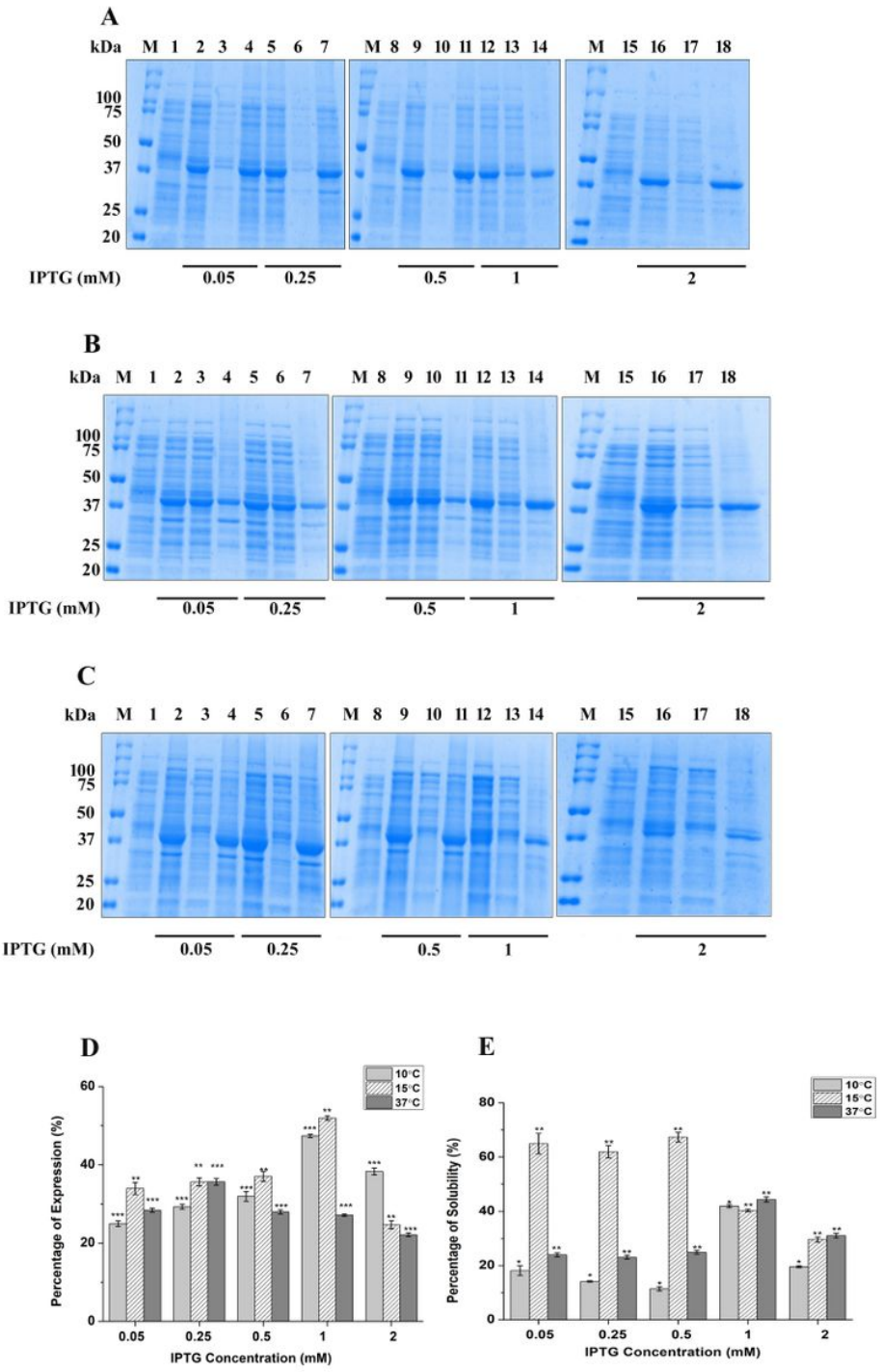

Figure 2

Expression and solubility analysis of pHis-TEV-IpaD in BL21 (DE3)

A) Expression level of IpaD protein at 10°C induced with different IPTG concentration. M : Molecular Weight Size marker; Lane 1: Total uninduced IpaD protein from BL21(DE3)(negative control); Lane 2: Total protein after IPTG induction (0.05mM); Lane 3: Soluble fraction after IPTG induction (0.05mM); Lane 4: Insoluble fraction after IPTG induction (0.05mM); Lane5: Total protein after IPTG induction (0.25mM); Lane 6: Soluble fraction after IPTG induction (0.25mM); Lane 7: Insoluble fraction after IPTG induction (0.25mM); Lane 8: Total uninduced IpaD protein from BL21(DE3) (negative control); Lane 9: Total protein after IPTG induction(0.5mM); Lane 10: Soluble fraction after IPTG induction(0.5mM); Lane 11: Insoluble fraction after IPTG induction (0.5mM); Lane 12: Total protein after IPTG induction(1 mM); Lane 13: Soluble fraction after IPTG induction(1mM); Lane 14: Insoluble fraction after IPTG induction(1 mM); Lane 15: Total uninduced IpaD protein from BL21(DE3) (negative control); Lane 16: Total protein after IPTG induction (2 mM); Lane 17: Soluble fraction after IPTG induction(2 mM); Lane 18: Insoluble fraction after IPTG induction (2 mM).

B) Expression level of IpaD protein at 15°C induced with different IPTG concentration. M : Molecular Weight Size marker; Lane 1: Total uninduced IpaD protein from BL21(DE3)(negative control); Lane 2: Total protein after IPTG induction (0.05mM); Lane 3: Soluble fraction after IPTG induction (0.05mM); Lane 4: Insoluble fraction after IPTG induction (0.05mM); Lane5: Total protein after IPTG induction (0.25mM); Lane 6: Soluble fraction after IPTG induction (0.25mM); Lane 7: Insoluble fraction after IPTG induction (0.25mM); Lane 8: Total uninduced IpaD protein from BL21(DE3) (negative control); Lane 9: Total protein after IPTG induction(0.5mM); Lane 10: Soluble fraction after IPTG induction(0.5mM); Lane 11: Insoluble fraction after IPTG induction (0.5mM); Lane 12: Total protein after IPTG induction (1 mM); Lane 13: Soluble fraction after IPTG induction(1mM); Lane 14: Insoluble fraction after IPTG induction (1 mM); Lane 15: Total uninduced IpaD protein from BL21(DE3) (negative control); Lane 16: Total protein after IPTG induction (2 mM); Lane 17: Soluble fraction after IPTG induction(2 mM); Lane 18: Insoluble fraction after IPTG induction(2 mM).

C) Expression level of IpaD protein at 37°C induced with different IPTG concentration. M : Molecular Weight Size marker; Lane 1: Total uninduced IpaD protein from BL21(DE3)(negative control); Lane 2: Total protein after IPTG induction (0.05mM); Lane 3: Soluble fraction after IPTG induction (0.05mM); Lane 4: Insoluble fraction after IPTG induction (0.05mM); Lane5: Total protein after IPTG induction (0.25mM); Lane 6: Soluble fraction after IPTG induction (0.25mM); Lane 7: Insoluble fraction after IPTG induction (0.25mM); Lane 8: Total uninduced IpaD protein from BL21(DE3) (negative control); Lane 9: Total protein after IPTG induction(0.5mM); Lane 10: Soluble fraction after IPTG induction(0.5mM); Lane 11: Insoluble fraction after IPTG induction (0.5mM); Lane 12: Total protein after IPTG induction(1 mM); Lane 13: Soluble fraction after IPTG induction(1mM); Lane 14: Insoluble fraction after IPTG induction(1 mM); Lane 15: Total uninduced IpaD protein from BL21(DE3) (negative control); Lane 16: Total protein after IPTG induction(2 mM); Lane 17: Soluble fraction after IPTG induction(2 mM); Lane 18: Insoluble fraction after IPTG induction (2 mM).

D) The effect of IPTG concentration on the expression level of the IpaD protein at different temperatures.

E) The effect of IPTG concentration on the percentage of solubility of the IpaD protein at different temperatures.

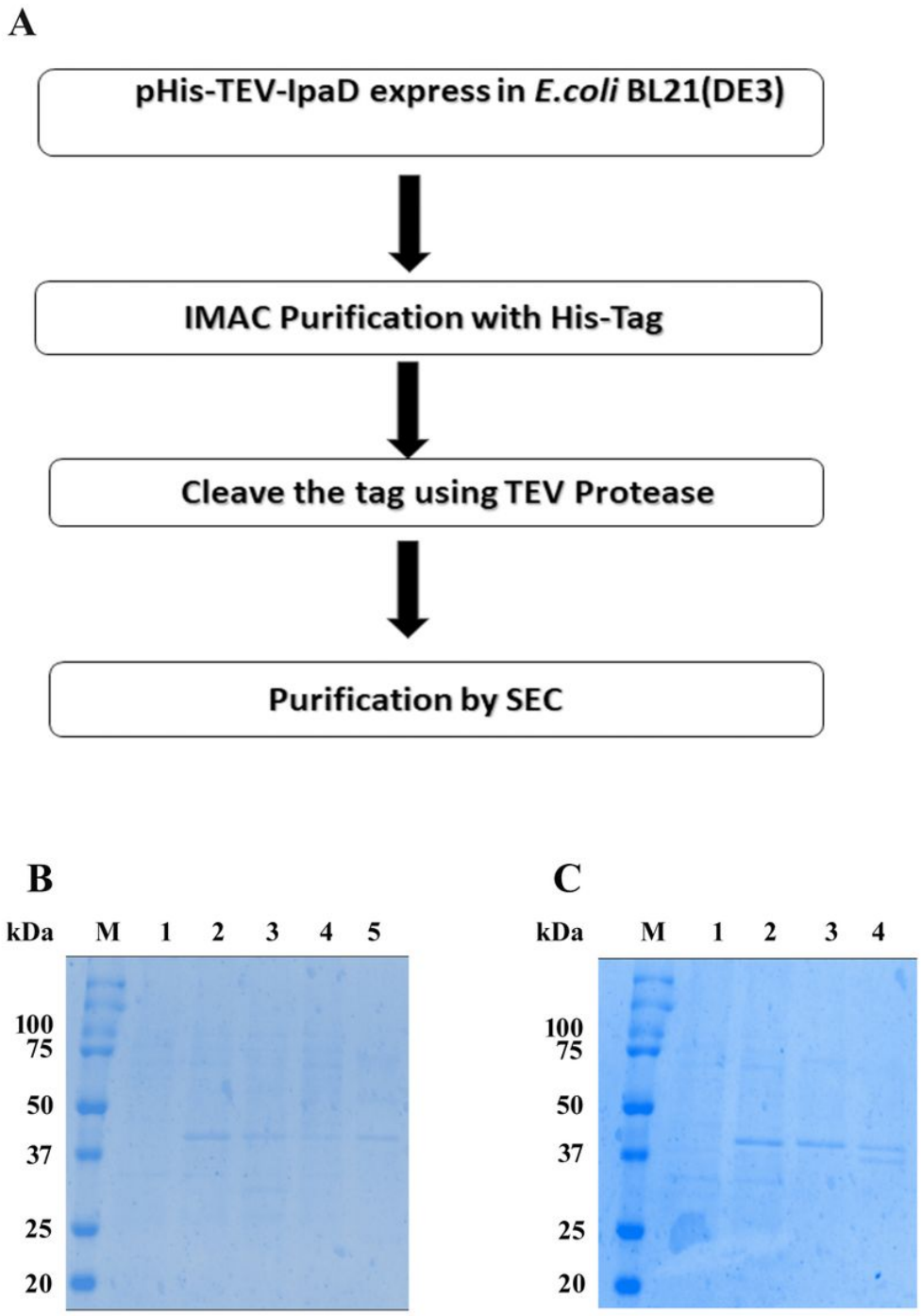

Figure 3

Purification of pHis-TEV–IpaD in BL21 (DE3)

A) Flow chart of purification Steps. IMAC, Immobilized metal ion chromatography; TEV, Tobacco etch virus; SEC, Size exclusion Chromatography

B) IMAC purified IpaD protein sample at different fractions. M: Molecular weight size marker; Lane 1: uninduced IpaD (negative Control); Lane 2: IPTG Induced IpaD in soluble fraction; Lane 3: Flow through fraction; Lane 4: wash fraction; Lane 5: Elution fraction.

C) pHis-TEV-IpaD purified from BL21 (DE3) M: Molecular weight size marker; Lane 1: Uninduced IpaD (negative Control); Lane 2: IPTG Induced IpaD in soluble fraction; Lane 3: IMAC Purified IpaD before cleave; Lane 4: IMAC purified IpaD after tev cleavage

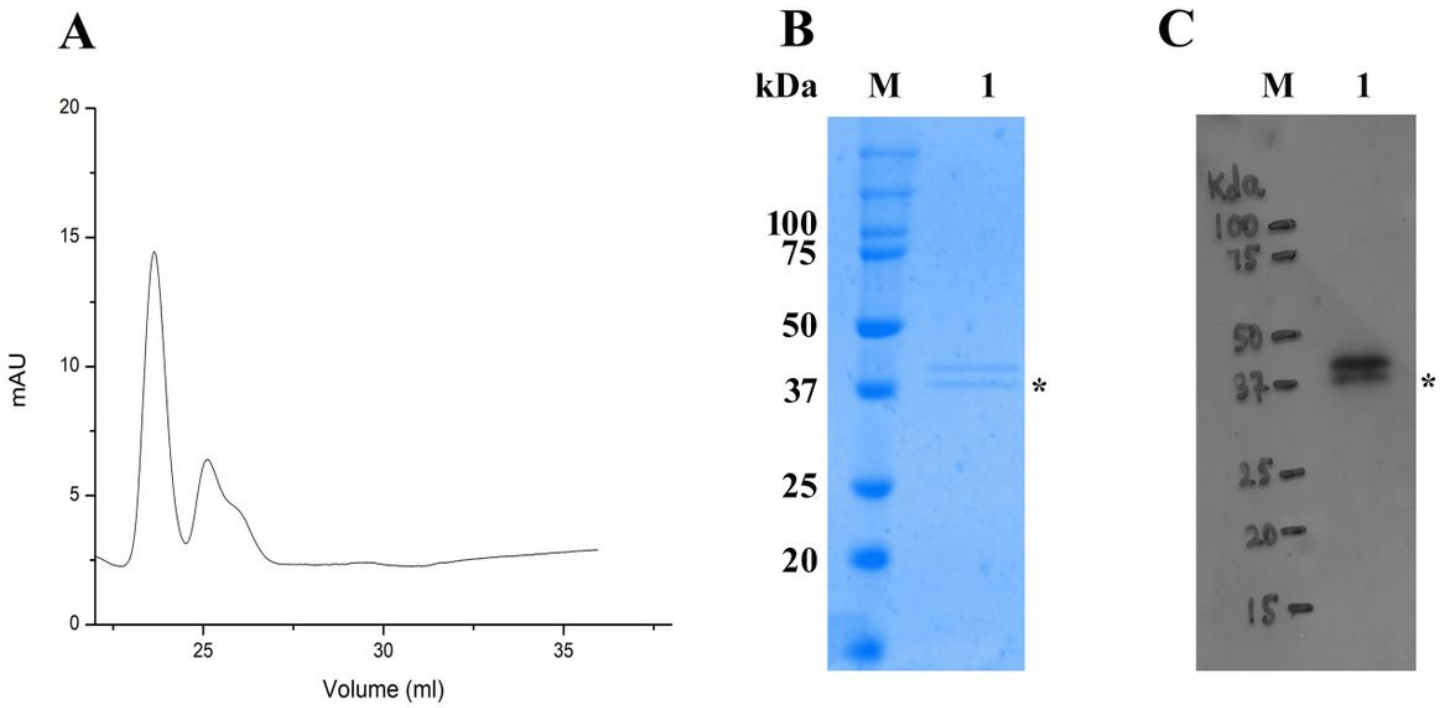

Figure 4

Determination of the purity of IpaD protein

A) Purified IpaD was analyzed by size exclusion chromatography to evaluate the purity. The x-axis shows the retention volume (ml) and the y-axis indicates the absorbance at 280 nm (arbitrary units, AU). The main peak of IpaD was observed between the retention volume 20-25ml

B) Purified fractions from (A) were analyzed using SDS-PAGE.M: Molecular Weight Size marker; Lane 1: Purified IpaD;

C) Western blot analysis to assess the purity of the IpaD protein; M: Molecular Weight Size marker; Lane 1: Purified IpaD whereas the asterisk (*) indicates truncations of the IpaD protein

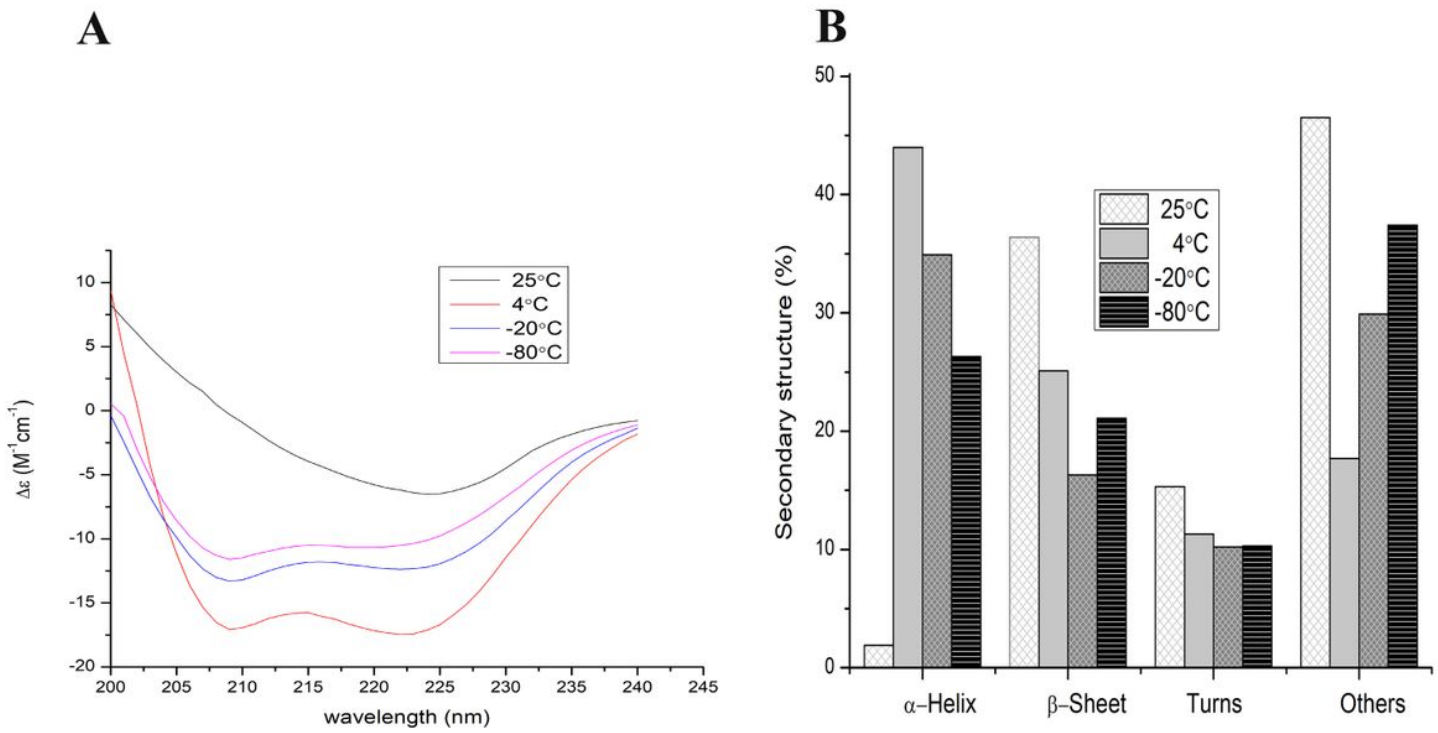

Figure 5

Analysis of the secondary structure of IpaD protein

A) Effect of the different storage temperature on secondary structure of purified IpaD protein determined by far UV CD spectroscopy. CD spectra have been plotted with wavelength (nm) against Δε ($M^{-1}cm^{-1}$).

B) The bar graph shows the percentage of secondary structures [α-helix, β-sheet, turn, and other (random coil)] of IpaD protein

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Authorchecklist.docx
- SupplementaryMaterial.docx