Biochemistry and Biophysics Reports 21 (2020) 100693

Contents lists available at ScienceDirect

Biochemistry and Biophysics Reports

journal homepage: www.elsevier.com/locate/bbrep

Structural investigation of APRs to improve the solubility of outer membrane protease (PgtE) from *Salmonella enterica* serotype typhi- A multi-constraint approach

Gopinath Samykannu[a,*], Princy Vijayababu[a], Christian Bharathi Antonyraj[b], Sundarabaalaji Narayanan[a,**]

[a] Structural Biology Laboratory, Department of Bioinformatics, Bharathiar University, Coimbatore, 641046, Tamil Nadu, India
[b] Centre for Bioinformatics, Pondicherry University, Pondicherry, India

ARTICLE INFO

Keywords:
Salmonella enteric serotype typhi
PgtE
APRs
Solubilization
CABS flex
Structural gatekeepers

ABSTRACT

Outer membrane proteins were playing a crucial role on the several functions controlled by cell membranes even though they are not naturally expressed at higher levels. In order to obtain biologically active protein, the denaturation of these inclusion bodies must be optimized using chaotropic agents. Hence, this study focuses on improving the yield of Outer Membrane Protease (PgtE) from *Salmonella enterica* serotype Typhi (*S*. Typhi) using chaotropes and additives. Denaturation methods were tried with various pH, detergents, and reducing agents were used to optimize the solubility of PgtE with biologically active form. Due to the aggregation, we failed to achieve the maximum yield of PgtE. Consequently, we predicted 9 Aggregation Prone Regions (APRs) in PgtE, which are mutated by known structural Gatekeepers. We calculated the Aggregation Index (AI) of PgtE with 10 mM of aspartic acid as an additive in optimized buffer. In addition, the mutations at specific positions within the protein structure can act as APRs suppressors without affecting protein stability with CABS flex dynamics. The multiple sequence analysis demonstrate that aspartic acid is appropriate denaturing additive for other Gram-negative pathogens of Omptin family.

1. Introduction

Membrane proteins play a significant role on the function controlled by cell membranes in both prokaryotic and eukaryotic organisms. They are regulating numerous biologically important functions such as transport of ions into and out of cells across the membrane, recognizing the immune system, energy transducers, mediating membrane adhesion and serving as a receptor for signal molecules [1–3]. Nearly 50% of human proteins are intrinsic or peripheral to cellular membranes [4]. Approximately 60% of membrane proteins were used as a drug targets in the pharmaceutical industries for several disease. Although, the importance of membrane proteins in cell function, extraordinary deficit of membrane protein structures is largely due to the insolubility of membrane proteins in aqueous solution [1,5–7]. Membrane proteins were consisting of more hydrophobic patches on their surfaces that enable the proteins to occupy the hydrophobic interior of a lipid bilayer in a membrane [8,9].The Outer Membrane (OM) of Gram-negative bacterial organism forms a protective permeability barrier around the cells, and serves as a molecular filter for hydrophilic substances. The channels were present in the OM, which leads to mediate the transport of nutrients and ions across the membrane into periplasm [10,11]. The PgtE from *Salmonella enteric* serotype Typhi (*S*. Typhi) is a member of omptin family and its molecular weight about 33.5 kDa of integral Outer Membrane protease that cleaves peptides and proteins preferentially between two consecutive basic amino acids [12]. In omptin family, the Gram-negative bacterial pathogens such as OmpT and OmpP of *E. coli*, *PgtE* of *S. typhimurium*, *Pla* of *Y. pestis* and *SopA* of *S. flexneri* were consists of Outer Membrane [12]. Among these proteins, the PgtE has been found to be involved in DNA excision repair mechanism with the breakdown of antimicrobial peptides in the urinary tract diseases, but its actual biological function remains to be elucidated [13]. Obtaining biologically active membrane protein is a challenging task because of their poor solubility and yield [14–19]. In a majority of cases, the collection of overexpressed protein forming a dense amorphous insoluble aggregates within the cell known as inclusion bodies [20,21]. Therefore, a cost-effective and downstream process is required for

* Corresponding author.
** Corresponding author.
E-mail address: gopinathsamykannu@gmail.com (G. Samykannu).

https://doi.org/10.1016/j.bbrep.2019.100693
Received 18 May 2019; Received in revised form 3 September 2019; Accepted 2 October 2019

isolation of intracellular inclusion bodies with higher solubility with good refolding capacity will desire their yield [22]. Size and density of the inclusion bodies depends upon their expression level and nature of protein. To prepare very pure inclusion bodies for subsequent characterization studies were obtained by washing with a buffer containing detergent and additives to remove the protein aggregation [23,24]. Due to the lack of several limitations in the bulk production of many potentially valuable therapeutic agents and membrane proteins were resembling and the PgtE are remains beyond the production [25].

Indeed, reducing of protein aggregation is the straightforward task to induce the active protein production [26]. The aggregation nature of a protein strongly determined by the intrinsic properties of its amino acids sequence and its mediated by short Aggregation Prone Regions (APRs) [27,28]. Over 80% of proteins consist of at least one APR in their primary sequence [29], but most often they form an integral part of the hydrophobic core of the protein [30,31], which leads to the lack of higher-order interactions. Fascinatingly, recent algorithm (AGGRESCAN) proposing that the aggregation of the protein is not only based the primary sequence, and other factors also responsible that properties such as experimental proteome data, aggregation propensity of bacterial proteins associated with their length, conformation, location, function and abundance of the polypeptide [32]. Aggregation is very difficult task to remove without affecting the stability and structure of protein. The reduction of intrinsic aggregation is usually achieved by introducing of aggregation breaking residues in a protein, called gatekeepers [33,34].

In general, the denaturation of proteins directly related to the number of APRs found within a macromolecule sequence. In this study, successful attempts have been made to increase the denaturation of *S.* Typhi PgtE by using aspartic acid as an additive, which has been found to reduce the number of APRs is replaced by introducing of structural gatekeeper in the loop regions of the macromolecule structure. These biologically active PgtE that can be used for structure functional studies in future.

2. Materials and methods

2.1. Cloning

The *pgtE* gene was isolated from the products obtained by PCR amplifications with the gene specific primers (Forward primer: **5′** GCCCCCATATGCTTAATATTCAAA ATTATTCCGC **3′** and Reverse primer: **5′**CGCCATGGTTATCCTTGCAGGAAGC**3′** containing *NdeI* and *Nco*I restriction sites, respectively**)** using *S.* Typhi *Ty21a* Genomic DNA as the template. Amplified gene product was digested with the restriction enzyme sites at *NdeI* and *Nco*I (New England Biolabs) and followed by the ligation into the plasmid vector pET30b(+) (Novagen) for over expression. The ligation mixture was transformed into the DH5α competent cells of *E. coli* by $CaCl_2$ heat-shock method and plated on LB-plate containing kanamycin (30 μg/ml) antibiotic resistance followed by the incubation at 37 °C for 12 h. The transformed colonies were screened by colony PCR using gene specific primers. Plasmid DNA was isolated and verified for the presence of insert using restriction digestion analysis with *NdeI* and *Nco*I sites.

2.2. Overexpression

The recombinant plasmids were transformed into the BL21 (DE3), C43 (DE3) and Rosetta (DE3) competent cells of *E. coli* were used as expression host strains. A single colony was selectively picked from the transformed colonies and subsequently transferred into a 50 mL of LB medium supplemented with kanamycin (30 μg/mL). The bacterial cells were allowed to grow inside an orbital shaker incubator at 160 rpm/min 37 °C. When the OD_{600nm} of each flask reached 0.4 to 0.6, the protein production was induced by the addition of IPTG with different concentrations (0.2, 0.4, 0.6, 0.8, 1.0, 1.5, and 2.0 mM) and grown at 37 °C for 4hrs. The temperature optimization of bacterial cell culture was performed at four different temperatures (20, 25, 30 and 37 °C) with 0.4 mM of IPTG and incubated for 3 and 6 h. For the time optimization of bacterial culture was induced with 0.4 mM of IPTG at 20 °C for 4, 8, 12, 16, 20, and 24 h respectively. The absorbance of all the samples from each experiment was measured at OD_{600nm} to monitor the bacterial growth. The maximum cell growth of bacterial cells achieved at OD_{600nm} and harvested at 10000 rpm for 10 min at 4 °C and followed by 12% SDS–PAGE analysis.

2.3. Extraction and purification of PgtE inclusion bodies

In order to, check the existence of PgtE from *S.* Typhi in pellet (inclusion bodies), and suspended in 50 mM of tris-HCl (pH 8.0) without any denaturant. The cells were disrupted by sonication with ultra-sonic waves generated by SONICS, 130 W using an optimized run time of 30 min with a pulse of 3 s on and 6 s off on ice. The crude inclusion bodies were obtained by the centrifugation at 10,000 g for 10 min at 4 °C. The presence of inclusion bodies in the pellet were confirmed by 12% SDS-PAGE and washed thrice with 10 mL of buffer A (50 mM tris - HCl (pH 8.0), 150 mM NaCl, 1% Triton X-100 and 2 M Urea). The extensive final wash of inclusion bodies was performed with buffer A without 1% Triton X-100 and 2 M urea followed by re-suspension and the solution kept on a shaking incubator at 37 °C for 15 min and followed by the centrifugation at 7000 rpm for 7 mins at 4 °C to remove soluble and contaminated proteins. The supernatant and pellet were obtained during every stages of the purification process were analyzed by 12% SDS PAGE.

2.4. Optimization of denaturation using various conditions

The insoluble pellet fractions of recombinant PgtE were undergone with numerous optimization steps such as using different concentrations of IPTG (0.2, 0.4 and 1 mM), varying the temperature (16, 20, 24,37 and 42 °C) and incubation time (30mins, 1, 2, 4, 8, 12, 16 and 24hrs), with various host cells ((BL21(DE3), C41(DE3), C43(DE3) and Rosetta (DE3)). Also, the insoluble pellets were tried with different media (Terrific Broth-TB, Luria-Bertani broth-LB, Minimal media, LB& TB-Auto induction media), pH (3, 4, 5, 8, 9, 10, 11 and 12), additive (Glycine) and different molar concentration of chaotropic agents (Urea and GdmHCl) were attempted to express the PgtE from *S.* Typhi.

2.5. Denaturation of PgtE with chaotropes (urea and GdmHCl)

The purified IBs were dissolved in 8 M urea with 50 mM of tris-HCl buffer at pH 8.0 and subjected to sonication for shorter duration (3 s). After leaving the suspension, the lysate kept for 3hrs, 6hrs and overnight at 37 °C and followed by centrifugation at 15,000 rpm for 90 min at 20 °C. The resultant supernatant and pellet were analyzed on 12% SDS-PAGE. Further denaturation carried out with the same method with the resultant insoluble PgtE pellet obtained. All these methods were used Chaotropes (Urea or GdmHCl) [35,36].

2.6. In silico identification of aggregation prone regions (APRs)

The primary sequence of *S.* Typhi PgtE comprising 312 amino acid residues were retrieved from UniProtKB (Q8Z4Y4) [37] and followed by the analysis of sequence by ProGene1.0 [38] for calculating molecular weight and theoretical Iso electric point (PI). AGGRESCAN server is a web-based application used to predict APRs in PgtE primary sequence. The aggregation-propensity value for each residue in the PgtE sequence was computed by averaging aggregation-propensity value per amino acid over a sliding window of a given length [32,39].

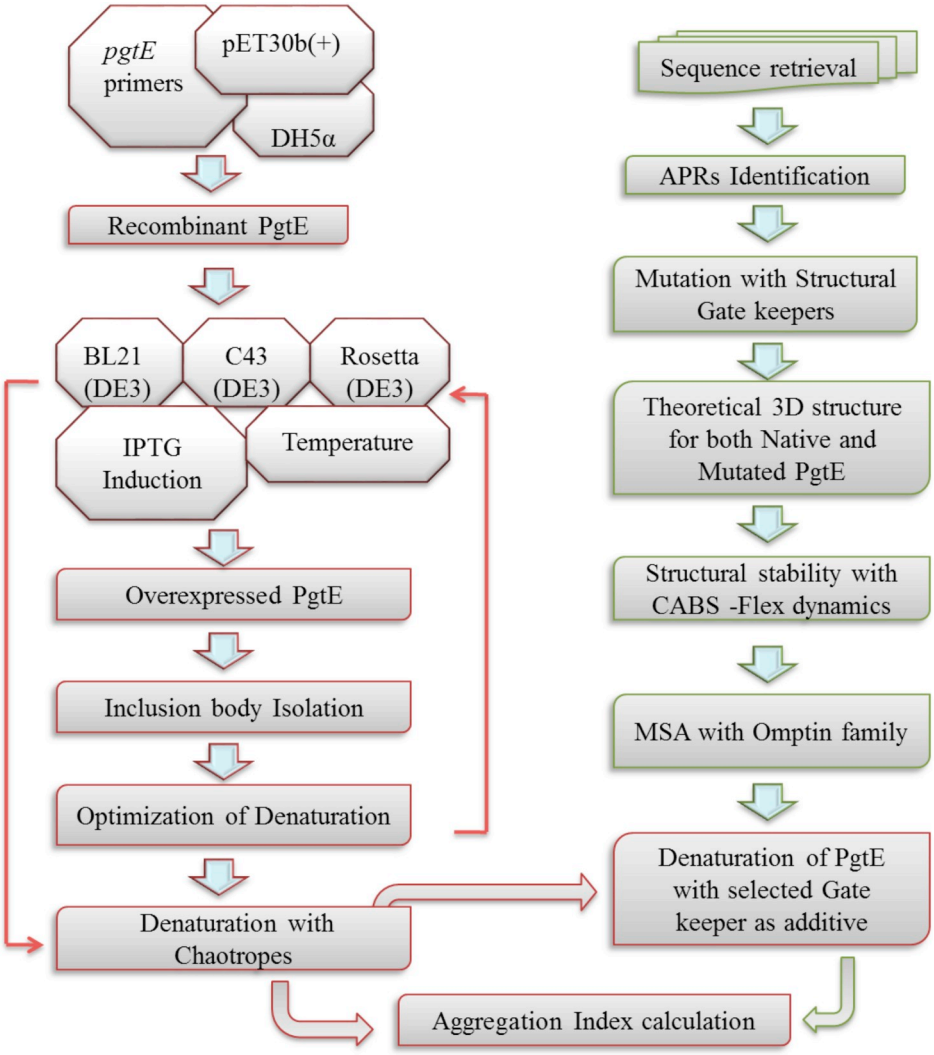

Fig. 1. *Overview - Denaturation of Outer Membrane Protease E from S.* Typhi.

2.7. Structural stability with additive mutants

The aggregation propensity of each single residue is calculated by AGGRESCAN server based on their relative position in the sequence. APRs were identified based on the average value of aggregation propensity and mutated with structural gatekeepers (P, R, K, D and E) at the hydrophobic residue in the centre of APRs. Therefore, in order to analyze the conservative sequences of APRs in protein sequence of outer membrane proteases, which belongs to the Gram-negative bacteria of Omptin family were retrieved and subjected for multiple sequence alignment.

Simultaneously, the homology structures of native and mutated proteins were determined by I-TASSER program [40] for the structural stability analysis. The best models were selected based on their computed C-score. The selected 3D structures were subjected to Modrefiner [41] for energy minimization and followed by the validation using the online version of PROCHECK program [42]. Theoretically, modeled native and mutated structures of PgtE were also subjected into CABS - Flex server to know the structural flexibility at 10 ns time scale [43].

2.8. Denaturation of PgtE from IBs with additives

Purified IBs were harvested and resuspended in 30 mL buffer containing 8 M of urea and further diluted into the buffer containing 50 mM tris-HCl (pH12.0), 100 mM NaCl, 2% Triton X-100, 5 mM DTT and 10 mM aspartic acid as additive increased the yield with moderate shaking at 37 °C for overnight. After the incubation, the solution containing PgtE was centrifuged at 15,000 rpm for 90 min at 20 °C. The separated supernatant and pellet were examined by SDS-PAGE.

2.9. Aggregation index (AI) – Calculation

The calculation of Aggregation Index done by UV spectrometer (T80 + UV/VIS spectrometer), which is useful for analyzing the various levels of aggregates present in protein solutions. The amount of aggregates was calculated using a simple equation that assigns an Aggregation Index (AI) value to a protein is AI = 100 X (Abs 340 nm/(Abs 280-Abs 340)). An overview of methodology used in our current study is shown in Fig. 1.

3. Results and discussion

3.1. Production of recombinant PgtE

The *pgtE* gene from *S.* Typhi was cloned in the multiple site of prokaryotic expression vector pET30b(+) between the restriction enzyme sites *NdeI* and *Nco*I. Double digestions were conducted in both amplified gene as well as in the plasmid vector for the ligation of targeted gene sequence about 900bp in to pET30b(+) as shown in Fig. 2.

Positive clones were selected from the products of colony-PCR using unambiguous primers of *pgtE* (mentioned methodology). The cloned

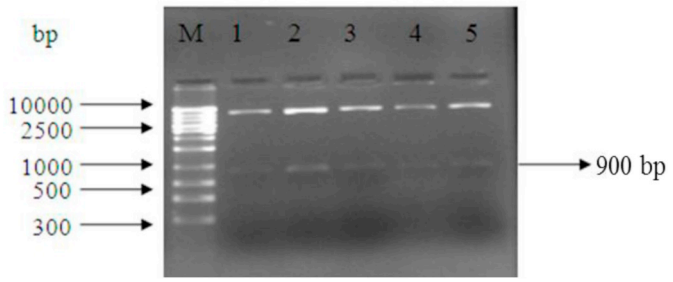

Fig. 2. Restriction digestion of pgtE. Two selected clones were digested with corresponding restriction enzymes NdeI and NcoI. Insert was released at the position of ~900 bp.

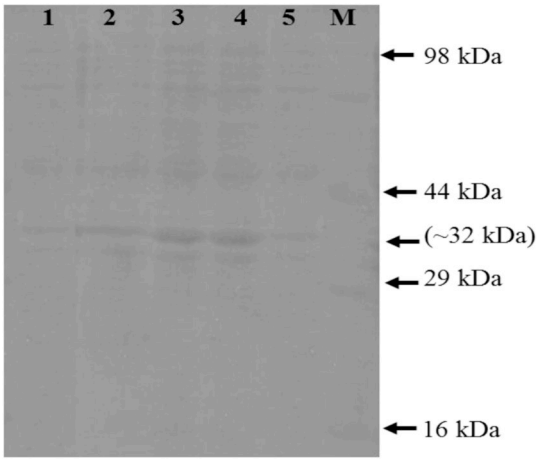

Fig. 3. Over expression of PgtE. Lane 1 and 2: pET30b (+) – pgtE - Uninduced samples (Clone #01 and 02), Lane 3 and 4: pET30b (+) - pgtE, Induced samples (Clone# 01 and 02), Lane 5: pET30b (+), Induced sample and Lane M – Molecular weight Marker.

plasmid was transformed into BL21 (DE3) competent cells as expression host of *E. coli* cells. Expression of cloned gene was under the control of a T7 universal promoter, which can be easily induced by 0.4 mM of isopropyl β-D-1-thiogalactopyranoside (IPTG) with an increased presence of a band showing an approximate molecular weight of 32 kD as shown in Fig. 3.

3.2. Effect of various agents on denaturation of PgtE

The rate and yield of the solubility process seem to be influenced by the conditions used, such as host strains, temperature, incubation period, pH, chaotropes, detergents and other additives. Even if there is no general procedure to increase the solubility of a protein suggest that in each case should be "custom-made". The procedure for the solubility and isolation of IBs consists of following three steps [44,45]. (i) over-expression of the selected protein in an appropriate host strain; (ii) isolation of IBs; (iii) Denaturation of IBs are the most critical steps in the procedure with trial-and-error strategies[44,46]. Based on the above said established procedures for our trial and error strategies for over expression of PgtE included so many conditions such as host, temperature, growth media and enhanced Induction were used to obtain the maximum yield of PgtE as shown in Table 1. The strain or genetic background of host organism is important for the expression of recombinant proteins. To bring the soluble form of recombinant PgtE, we tried different host cells such as C41(DE3), C43(DE3), Rosetta and BL21(DE3) to improve yield of PgtE. None of the host cells showed any improvement on the denaturation of PgtE except BL21(DE3) competent cells of *E. coli* as a host were used to express the PgtE in subsequent steps. Temperature is one of the important parameter that affects the expression of recombinant protein [47]. The expression level of PgtE was found to be higher at 20 °C, where as in the low temperature increases the activity of expressed protein owing to slower rate of protein synthesis [48]. The induction temperature of targeted gene optimized in the range of 16–23 °C [49]. Suggested that cells growth of PgtE was found to be maximum at 20 °C for 6 hrs at 20 °C. Higher temperature protein essentially emerged as higher levels of insoluble aggregates.

The composition and richness of media have more significant effect on protein expression and solubility of desired protein [50]. PgtE construct was tried in different media such as LB, minimal media, TB, LB-AIM and 2XYT for the bacterial cell growth. The cells grown in LB and TB media showing higher level of cell mass than the minimal media. However, they were not helpful in lowering the protein aggregation of targeted gene. The minimal media consists of glucose, other salt elements such as magnesium, nitrogen, phosphorus, and sulfur to allow the bacteria to grow and synthesizing of protein. As a result, the minimal media was chosen as a suitable one for the subsequent trails in future and the concentration of IPTG needs to be optimized for every protein under induced conditions [51]. In this study, different concentrations (0.2–1.4 mM) of IPTG were used for increasing the expression level and solubility of targeted protein. The lower concentration of inducer will increase the yield of protein [52,53]. The expression level of PgtE is maximum at 0.4 mM of IPTG used and finalized as the optimal concentrations for the expression of PgtE.

3.3. Denaturation of PgtE with chaotropes

Chaotropic mechanism suggests that GdmHCl acts as a chaotrope, altering the hydrogen-bond structure of water, hence lessening the effects of hydrophobic interactions, and promoting solvation of hydrophobic groups and pH deviation from the physiological conditions leads to either protonation of the carboxylic groups (decrease in pH) or deprotonation of the amino groups (increase in pH) which changed local charges and subsequently resulted in repulsion that destabilize the protein [54–57].

Denaturation of PgtE was tried by chaotropic agents such as urea and guanidine hydrochloride (GdmHCl) to increase the yield [54,55]. The concentration of urea (2 M – 8 M) or GdmHCl (2 M – 6 M) were gradually increased in the buffer conditions to achieve the denaturation of IBs in different pH as shown in Table 2 and Supplementary Fig. 1. From these results, 8 M urea has been found to be the suitable concentrations to increase the solubility of PgtE, which used for further purification steps.

Induced and pre induced (uninduced) samples were also analyzed

Table 1
Various factors affects the overexpression of PgtE for maximum yield.

Expression Host	Temperature(°C)	Growth media	IPTG Induction (mM)	Overexpression Observations
C41(DE3), C43(DE3), Rosetta	37, 20	LB, MM	0.4	X
BL21(DE3)	37,20,16,4	LB,MM	0.2	X
		LB	0.4	X
	20	MM		**Good**
			0.6,0.8,1.0,1.2,1.4	Not much variation
		LB-AIM	–	X
		MM-AIM		X

Table 2
Denaturation of PgtE with various buffer composition.

Urea (M)	tris- HCl Buffer (pH)	NaCl (mM)	Reducing agent (β –ME) (mM)	Aggregation-suppressor Agent(Glycine) (mM)
4 M − 8 M	pH:8.0	100 mM	–	–
8 M	pH:3.0–6.0 & 9.0–12.0	100 mM	–	–
8 M	pH: 9.0	100 & 150 mM	–	–
8 M	pH:8.0 & 8.5	150 mM		10,20,30,40,50 mM
8 M	pH:8.0	150 & 200 mM	6,7 & 8 mM	10,20 & 100 mM
8 M	pH:8.0 & 10.0	300 & 500 mM	–	50 mM
8 M	pH:10.0	150 mM	–	10,20 & 30 mM
8 M	pH:8.0 & 12.0	150 mM	1 mM	–
9 M	pH:8.0	150 & 200 mM	–	50 &100 mM
10 M	pH:8.0	150 mM	–	–

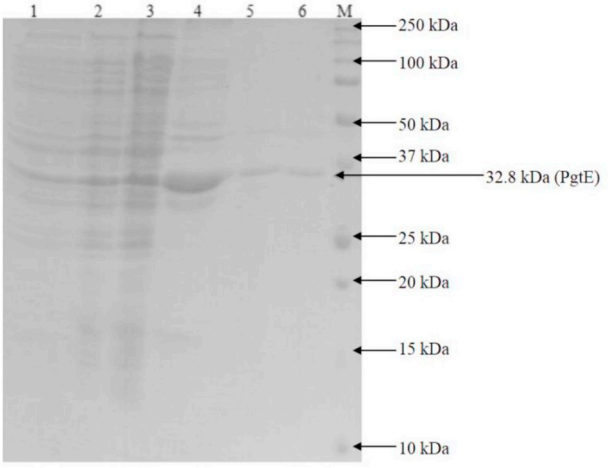

Fig. 4. PgtE one and two-step denaturation with Chaotropes (GdmHCl and Urea): Lane 1: Whole cell pellet, lane 2: Whole cell lysate, lane 3: Lysate supernatant, lane 4: PgtE -Inclusion bodies, lane 5: PgtE unfolded (denaturation) in supernatant (after 3hrs incubation at 37 °C), lane 6: PgtE unfolded in supernatant (after 12hrs incubation at 37 °C) and lane M: Molecular weight marker.

and various buffer composition were used (pH and tris-HCl) were shown in Supplementary Table 1. However, the protein was still found in pellet (shown in Fig. 4).

3.4. In-silico identification of APRs

The prediction of APRs in PgtE from *S.* Typhi were identified based on their primary sequence of a protein consist of 312 AA including signal peptide (20AA). The N-terminal signal peptide sequences were excluded during the calculation of AI to reduce the protein aggregation of a protein through *in silico* methods. Finally, the primary sequences of PgtE were analyzed by ProGene 1.0 for the calculation of molecular weight and isoelectric point about 32.8 kDa and 5.9 respectively. The results of AGGRESCAN server showing the presence of 9 APRs in PgtE from *S.* Typhi as follows: 3 SVTTSLSVGVL 23, 62 YSFMTLD 68, 116 VKGWL120, 162 GVRGI 166, 176 YIGLAG 181, 187 ECNVLFKYSD 198, 225 YGASIDAGYYITSNA237, 241 IFAEFAY 247 and 284 VTAGLQYRF 292 (shown in Fig. 5). In order to eliminate the APRs in PgtE from *S.* Typhi carried out by the replacing of amino acids at the centralized position of each APRs with the aggregation gatekeepers (P, R, K, D and E) through point mutations shown in Supplementary Fig. 2. In this connection, the following amino acids were replaced by the point mutation, which leads to decrease the protein aggregation as follows: L18D, M65D, G118D, R164D, G178P, L193D, G232D, E244P and L288D. The parameters used for the calculation of aggregation propensity with mutated residues were calculated individually for each mutations shown in Table 3. As the result of these point mutations were clearly shows that the hotspots become null hot spot as shown in Fig. 6A and B. From these results, aspartic acid has been found to be the best structural gatekeeper for PgtE from *S.* Typhi.

To know the conserved APRs and mutation in other Gram-negative

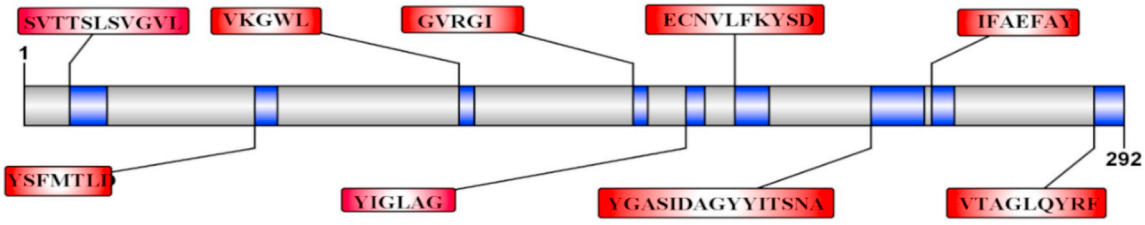

Fig. 5. Identified APRs in PgtE without signal peptide. Blue color box represents the position of APRs in primary sequence.

Table 3
Aggregation-propensity value for point mutation with structural gatekeepers.

	PgtE	M/P	M/R	M/K	M/D	M/E	M/DP
Aggregation-propensity value Sequence Average (a3vSA)	−0.096	−0.109	−0.137	−0.128	−0.156	−0.143	−0.15
Number of Hot Spots (nHS)	9	4	6	6	2	6	0
Normalized nHS for 100 residues (NnHS)	3.082	1.37	2.055	2.055	0.685	2.055	0
Area of the profile Above Threshold (AAT)	21.15	16.063	13.264	14.904	10.722	12.403	11.216
Total Hot Spot Area (THSA)	16.741	5.902	6.957	8.924	2.681	6.364	0
Total Area (TA)	−23.997	−27.701	−35.804	−33.041	−41.135	−37.343	−39.507
Area of the profile Above Threshold per residue (AATr)	0.072	0.055	0.045	0.051	0.037	0.042	0.038
Total Hot Spot Area per residue (THSAr)	0.057	0.02	0.024	0.031	0.009	0.022	0

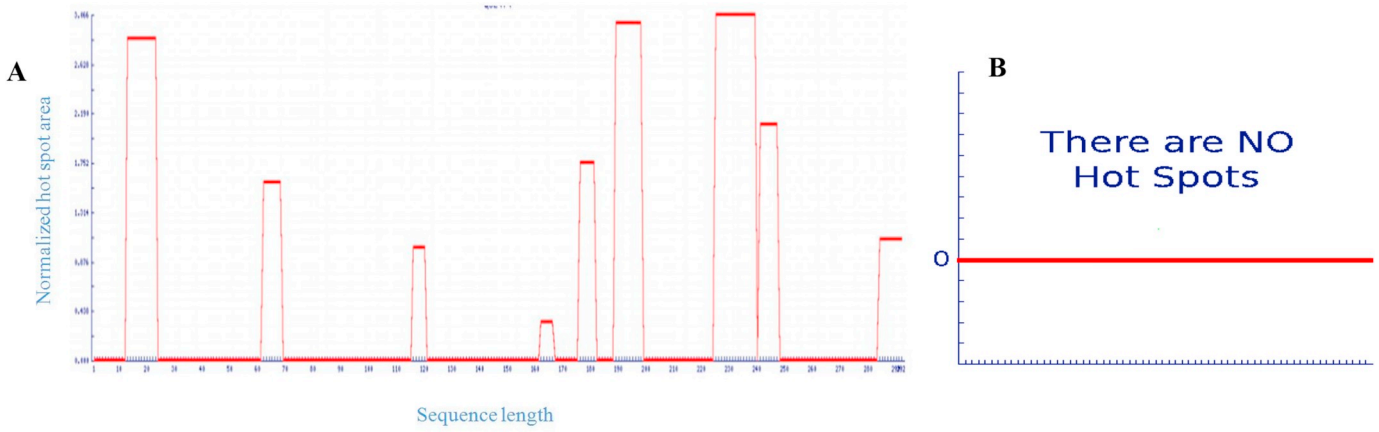

Fig. 6. A) Predicted total area of aggregation (total number of APRs) 6. B) Null APRs because of the dominancy mutation of aspartic acid and proline.

bacterial outer membrane protease from Omptin family (PgtE and OmpP of *E. coli*, PgtE of *S. typhimurium*, Pla of *Y. pestis* and SopA of *S. flexneri*) were retrieved from UniprotKB database. Multiple sequence alignment was carried out through BlastAlign and viewed by ESPript. The consequential alignment interprets mutation G118D, G178P, G232D, E244P and L288D are found in the conserved regions, which were shown in Fig. 7.

3.5. Structural stability with additive mutants

With the nonexistence of experimental structure of PgtE from *S.* Typhi, using I-TASSER online server built theoretical 3D structure and the validations of modeled structures were analyzed by Ramachandran plot through PROCHECK program. The hydrophobic core of the protein structure maintained by the necessary point mutations are to be done only at the loop regions of APRs [31]. Similarly, in the present study

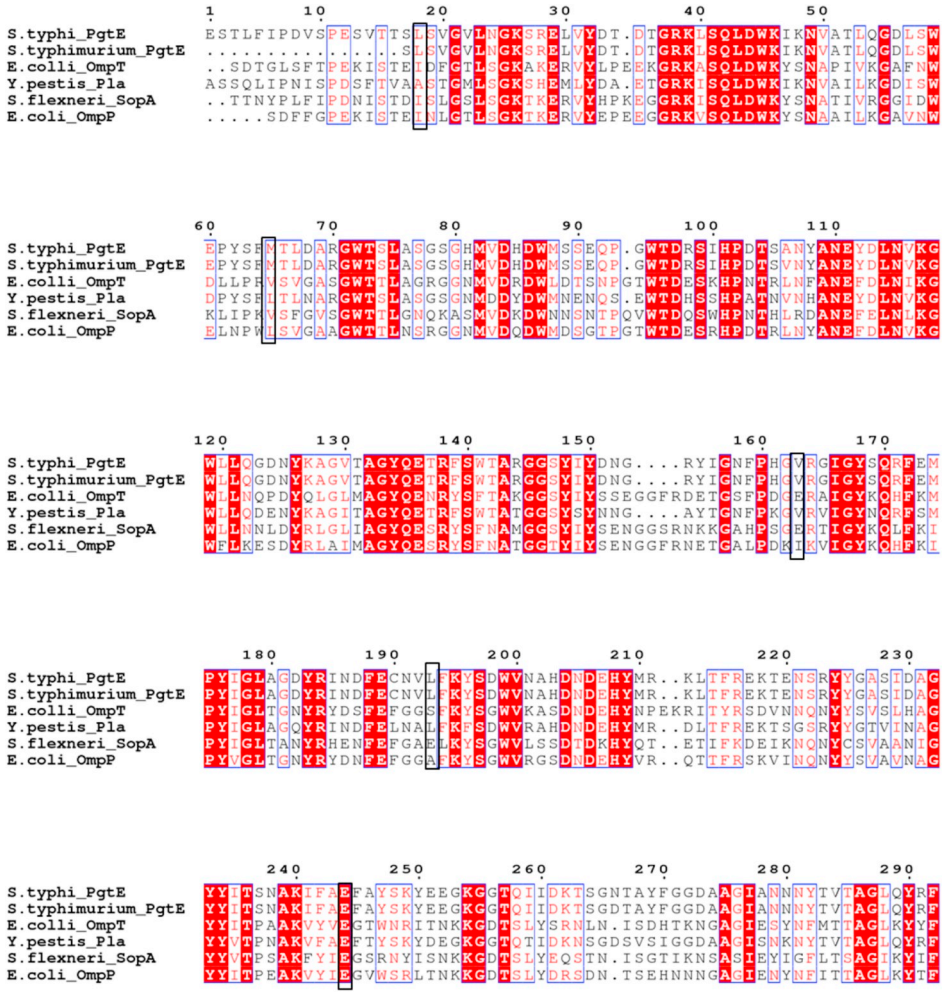

Fig. 7. Multiple sequence alignment of Gram-negative bacterial outer membrane protease from Omptin family. ESPript flat figure shows the identical and similar residues were differentiated in red and blue colour boxes. Red colour box and text shows the conserved and semi conserved regions respectively. Black colour highlighted box represents the (PgtE) mutated APRs at centralized position.

Table 4
CABS model of average cluster RMSD of both native and mutated PgtE.

Cluster #		1	2	3	4	5	6	7	8	9	10	11
Average cluster RMSD	Native	1.6	1.7	1.7	1.7	1.6	1.8	1.6	1.8	1.7	1.8	1.8
	Mutate	1.6	1.9	1.7	1.7	1.9	1.5	1.5	1.8	1.7	1.8	1.7

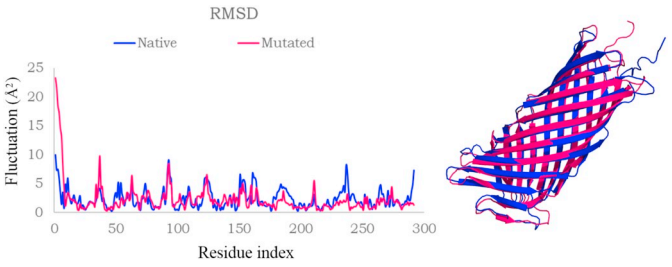

Fig. 8. RMSD fluctuation profile and secondary structure representation of both native and mutated PgtE.

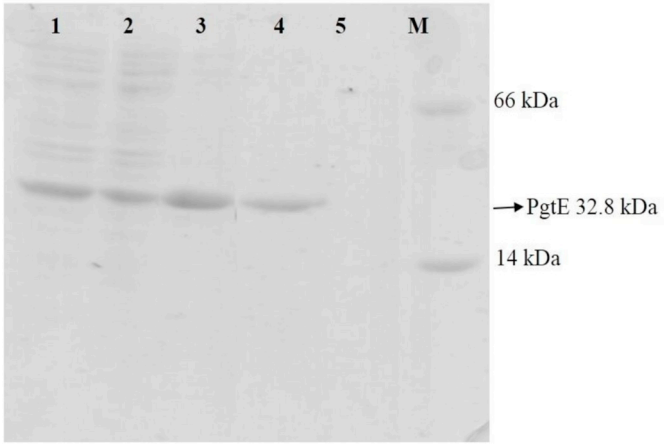

Fig. 9. Denaturation of PgtE using 8 M urea with aspartic acid. Lane 1: Whole cell lysate, lane 2: Lysate supernatant, lane3: PgtE Inclusion bodies, lane 4: PgtE unfolded in supernatant with aspartic acid, lane 5: PgtE unfolded in pellet and lane M − Molecular weight marker.

one of the mutations R164D is alone located in the loop region, whereas all other mutations were being exist in the β-sheets of the protein structure. Therefore, structural stability of native and mutated form of PgtE are analyzed.

CABS-flex server confirmed the MD simulations of PgtE in both native and mutated form at 10-ns time scale with all-atoms excluding water molecules with the force fields. The resulting trajectory of native and mutated form of PgtE were analyzed and clustered. About 11 PgtE models were generated using CABS model and the average cluster RMSD of both native and mutated structures were compared and shown in Table 4. The highest fluctuation of mutated PgtE about ~1.23 Å^2 of 162 GVRGI 166 with the mutation of R164D. The least fluctuation observed with the mutation of L288D in the corresponding APR (284 VTAGLQYRF 292) compared to the native form of PgtE (~0.16 Å^2). The fluctuation profiles of both native and mutated form of PgtE were shown in Fig. 8.

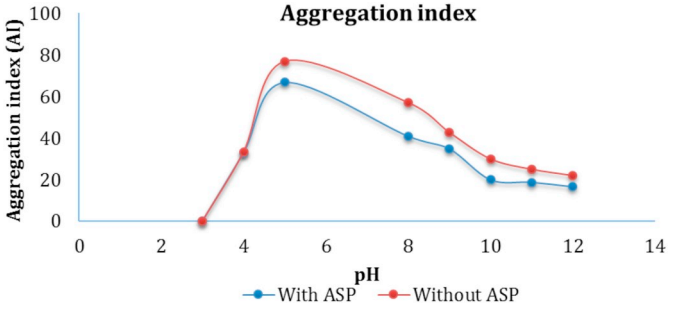

Fig. 10. PgtE-Aggregation Index were compared with presence and absence of aspartic acid in different pH.

3.6. Denaturation of PgtE from IBs

An aspartic acid is one of the structural gatekeepers were introduced as an additive in buffer composition (50 mM tris-HCl (pH12.0),100 mM NaCl, 8 M urea, 2% Triton X-100, 5 mM DTT and 10 mM aspartic acid), which is used for the denaturation of PgtE. Finally, the denaturation of PgtE was standardized with 8 M urea and 10 mM aspartic acid, which are evidenced that urea is a strong denaturing agent and aspartic acid as a good additive for the ability to increase the yield of PgtE than GdmHCl and other additives (Fig. 9). This standardized buffer could be used to further experiments for the solubilization (refolding) of members from omptin family.

3.7. Aggregation index (AI)

Protein aggregation index in unfolded form or solubility of protein was determined in solution by measuring of absorbance in the visible range of A_{280} andA_{340}. From the values of AI (Table 5) shows that the aggregation occurs in unfolded form of PgtE with a steady propensity towards the solubility with the different denaturants, pH and additives (Supplementary Table 1) are included in the buffer. The AI values of PgtE is above 20 suggested that the presence of significant levels of aggregation in the solution. At the point of extreme pH (12) [56,57], our AI results of PgtE is gradually decreases to 16.6 measured by the absorbance at A_{340} and A_{280} with the buffer contains 10 mM of aspartic acid as an additive shown in Fig. 10 and Supplementary Table 2. This study will clearly show that the aspartic acid will act as in aggregation reducing additive in *S.* Typhi PgtE.

4. Conclusion

The PgtE from *S.* Typhi gene was cloned into pET30b (+) and overexpressed under the control of T7 promoter using BL21 (DE3) of *E.coli* bacterial system. Our attempts to increase the denaturation of PgtE failed, due to the poor yield. Consequently, 9 APRs were identified by AGGRESCAN server and those regions were replaced by respective amino acids by the successful point mutations of L18D, M65D, G118D, R164D, G178P, L193D, G232D, E244P and L288D.Since the mutation aspartic acid (a charged residue) is a reminiscent of aggregation gatekeeper in most of the APRs and used as an additive for increasing the yield of PgtE. On the other hand, 10ns CABS-flex dynamics revealed that the mutations does not shows structural fluctuations with the RMSD of 1.48 Å between native and mutated structures. Results of multiple sequence analysis demonstrated that, the Gram-negative

Table 5
AI of unfolded sample showed that there were detectable aggregates in the solution.

Unfolded sample with Buffer Composition	A_{280}	A_{340}	AI = 100 X (Abs 340 nm/(Abs 280-Abs 340))
8 M Urea, 50 mM tris-HCl (pH12.0),100 mM NaCl, 2% Triton X-100, 10 mM Glycine and 5 mM DTT	0.009	0.002	28
8 M Urea, 50 mM tris-HCl (pH12.0), 100 mM NaCl, 2% Triton X-100, 10 mM Aspartic acid and 5 mM DTT	0.007	0.001	16.6

bacterial pathogens of Omptin family including OmpT and OmpP of *E. coli*, PgtE of *S. typhimurium*, Pla of *Y. pestis* and SopA of *S. flexneri* showing identical positions 6 out of 9 APRs in each protein. This may be appropriate to the other Outer Membrane Protease of Omptin family. To the best of our knowledge, this is the first report on the usage of aspartic acid for denaturation of PgtE from *S.* Typhi.

Acknowledgements

The authors thank to I. Karthik, CAS in Crystallography and Biophysics, University of Madras, Chennai-India., for valuable suggestion in the production of the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bbrep.2019.100693.

References

[1] E. Wallin, G.V. Heijne, Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms, Protein Sci. 7 (4) (2008) 1029–1038.
[2] G. Samykannu, P. Perumal, S. Narayanan, Computational and experimental characterization of integral membrane proteins at molecular level, in: InA.R. Rao (Ed.), Biotechnology: Bioinformatics and Computational Biology, stadiumpress LLC, USA, 2014, pp. 225–258.
[3] G. Samykannu, P. Vijayababu, J. Natarajan, Substrate specificities in Salmonella typhi outer membrane protease (PgtE) from Omptin family – an in silico proteomic approach, Informatics in Medicine Unlocked 12 (2018) 6–13.
[4] C. Sudhahar, R. Haney, Y. Xue, R. Stahelin, Cellular membranes and lipid-binding domains as attractive targets for drug development, Curr. Drug Targets 9 (8) (2008) 603–613.
[5] P. Vijayababu, G. Samykannu, C.B. Antonyraj, J. Thomas, S. Narayanan, S.I. Basheer Ahamed, S. Piramanayagam, Patulin interference with ATP binding cassette transferring auto inducer −2 in Salmonella typhi and biofilm inhibition via quorum sensing, Informatics in Medicine Unlocked 11 (2018) 9–14.
[6] M.A. Yıldırım, K.-I. Goh, M.E. Cusick, A.-L. Barabási, M. Vidal, Drug—target network, Nat. Biotechnol. 25 (10) (2007) 1119–1126.
[7] P. Vijayababu, G. Samykannu, C.B. Antonyraj, S. Narayanan, S.I. Basheer Ahamed, P. Perumal, S. Piramanayagam, B-cell and T-cell epitope identification with stability analysis of AI-2 import ATP-binding cassette LsrA from S. typhi in silico approach, Microb. Pathog. 123 (2018) 487–495.
[8] P. Vijayababu, G. Samykannu, S. Narayanan, Structural insights of cell-cell communication autoinducer-2 kinase (LsrK) model from Salmonella typhi, International Journal of Current Research 9 (05) (2017) 51169–51172.
[9] J.-H. Lin, A. Baumgaertner, Molecular dynamics simulations of hydrophobic and amphiphatic proteins interacting with a lipid bilayer membrane, Comput. Theor. Polym. Sci. 10 (1–2) (2000) 97–102.
[10] H. Nikaido, Outer membrane permeability of gram-negative bacteria in their interaction with the outside world, The Cell Membrane (1984) 7–17.
[11] G. Samykannu, P. Vijayababu, C.B. Antonyraj, P. Perumal, S. Narayanan, S.I. Basheer Ahamed, J. Natarajan, In silico characterization of B cell and T cell epitopes for subunit vaccine design of Salmonella typhi PgtE: a molecular dynamics simulation approach, J. Comput. Biol. 26 (2) (2019) 105–116.
[12] L. Vandeputte-Rutten, R.A. Kramer, J. Kroon, N. Dekker, M.R. Egmond, P. Gros, Crystal structure of the outer membrane protease OmpT from Escherichia coli suggests a novel catalytic site, EMBO J. 20 (18) (2001) 5033–5039.
[13] C.-Y. Hui, Y. Guo, Q.-S. He, L. Peng, S.-C. Wu, H. Cao, S.-H. Huang, Escherichia coli outer membrane protease OmpT confers resistance to urinary cationic peptides, Microbiol. Immunol. 54 (8) (2010) 452–459.
[14] G. Rabbani, M.H. Baig, K. Ahmad, I. Choi, Protein-protein interactions and their role in various diseases and their prediction techniques, Curr. Protein Pept. Sci. 19 (10) (2018) 948–957.
[15] S.K. Haq, G. Rabbani, E. Ahmad, S.M. Atif, R.H. Khan, Protease inhibitors: a panacea? J. Biochem. Mol. Toxicol. 24 (4) (2010) 270–277.
[16] M.V. Khan, M. Ishtikhar, G. Rabbani, M. Zaman, A.S. Abdelhameed, R.H. Khan, Polyols (Glycerol and Ethylene glycol) mediated amorphous aggregate inhibition and secondary structure restoration of metalloproteinase-conalbumin (ovotransferrin), Int. J. Biol. Macromol. 94 (2017) 290–300.
[17] M.V. Khan, G. Rabbani, E. Ahmad, R.H. Khan, Fluoroalcohols-induced modulation and amyloid formation in conalbumin, Int. J. Biol. Macromol. 70 (2014) 606–614.
[18] A. Singh, V. Upadhyay, A.K. Upadhyay, S.M. Singh, A.K. Panda, Protein recovery from inclusion bodies of Escherichia coli using mild solubilization process, Microb. Cell Factories 14 (2015) 41.
[19] G. Samykannu, P. Vijayababu, C.B. Antonyraj, S. Narayanan, S.I. Basheer Ahamed, Investigations of binding mode insight in Salmonella typhi type-III secretion system tip protein (SipD): a molecular docking and MD simulation study, Informatics in Medicine Unlocked 9 (2017) 166–172.
[20] D. Williams, R. Van Frank, W. Muth, J. Burnett, Cytoplasmic inclusion bodies in Escherichia coli producing biosynthetic human insulin proteins, Science 215 (4533) (1982) 687–689.
[21] R.G. Schoner, L.F. Ellis, B.E. Schoner, Isolation and purification of protein granules from Escherichia coli cells overproducing bovine growth hormone, Bio Technology 3 (2) (1985) 151–154.
[22] H.L. Cash, C.V. Whitham, L.V. Hooper, Refolding, purification, and characterization of human and murine RegIII proteins expressed in Escherichia coli, Protein Expr. Purif. 48 (1) (2006) 151–159.
[23] C.A. Valente, G.A. Monteiro, J.M.S. Cabral, M. Fevereiro, D.M.F. Prazeres, Optimization of the primary recovery of human interferon α2b from Escherichia coli inclusion bodies, Protein Expr. Purif. 45 (1) (2006) 226–234.
[24] W. Swietnicki, Folding aggregated proteins into functionally active forms, Curr. Opin. Biotechnol. 17 (4) (2006) 367–372.
[25] H. Li, X. Lin, S. Wang, X. Peng, Identification and antibody-therapeutic targeting of chloramphenicol-resistant outer membrane proteins in Escherichia coli, J. Proteome Res. 6 (9) (2007) 3628–3636.
[26] A. De Marco, L. Vigh, S. Diamant, P. Goloubinoff, Native folding of aggregation-prone recombinant proteins in Escherichia coli by osmolytes, plasmid- or benzyl alcohol–overexpressed molecular chaperones, Cell Stress & Chaperones 10 (4) (2005) 329–339.
[27] V. Castillo, R. Graña-Montes, R. Sabate, S. Ventura, Prediction of the aggregation propensity of proteins from the primary sequence: aggregation properties of proteomes, Biotechnol. J. 6 (6) (2011) 674–685.
[28] J. Beerten, J. Schymkowitz, F. Rousseau, Aggregation prone regions and gatekeeping residues in protein sequences, Curr. Top. Med. Chem. 12 (22) (2013) 2470–2478.
[29] F. Rousseau, J. Schymkowitz, L. Serrano, Protein aggregation and amyloidosis: confusion of the kinds? Curr. Opin. Struct. Biol. 16 (1) (2006) 118–126.
[30] G. De Baets, J. Van Durme, F. Rousseau, J. Schymkowitz, A genome-wide sequence–structure analysis suggests aggregation gatekeepers constitute an evolutionary constrained functional class, J. Mol. Biol. 426 (12) (2014) 2405–2412.
[31] A. Ganesan, A. Siekierska, J. Beerten, M. Brams, J. Van Durme, G. De Baets, ... J. Schymkowitz, Structural hot spots for the solubility of globular proteins, Nat. Commun. 7 (2016) 10816.
[32] O. Conchillo-Solé, N.S. de Groot, F.X. Avilés, J. Vendrell, X. Daura, S. Ventura, AGGRESCAN: a server for the prediction and evaluation of “hot spots” of aggregation in polypeptides, BMC Bioinf. 8 (1) (2007) 65.
[33] M.P. Repasky, M. Shelley, R.A. Friesner, Flexible ligand docking with glide, Current Protocols in Bioinformatics 18 (1) (2007) 1–8.
[34] E. Monsellier, F. Chiti, Prevention of amyloid-like aggregation as a driving force of protein evolution, EMBO Rep. 8 (8) (2007) 737–742.
[35] G. Rabbani, E. Ahmad, N. Zaidi, S. Fatima, R.H. Khan, pH-induced molten globule state of rhizopus niveus lipase is more resistant against thermal and chemical denaturation than its native state, Cell Biochem. Biophys. 62 (3) (2012) 487–499.
[36] G. Rabbani, I. Choi, Roles of osmolytes in protein folding and aggregation in cells and their biotechnological applications, Int. J. Biol. Macromol. 109 (2018) 483–491.
[37] R. Apweiler, A. Bairoch, C.H. Wu, Protein sequence databases, Curr. Opin. Chem. Biol. 8 (1) (2004) 76–80.
[38] P. Vijayababu, G. Samykannu, S. Narayanan, ProGene 1.0-an in silico tool for protein-gene analysis, Int. J. Life Sci. Sci. Res. 3 (3) (2017).
[39] M.V. Khan, G. Rabbani, M. Ishtikhar, S. Khan, G. Saini, R.H. Khan, Non-fluorinated cosolvents: a potent amorphous aggregate inducer of metalloproteinase-conalbumin (ovotransferrin), Int. J. Biol. Macromol. 78 (2015) 417–428.
[40] Y. Zhang, I-TASSER server for protein 3D structure prediction, BMC Bioinf. 9 (1) (2008) 40.
[41] D. Xu, Y. Zhang, Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization, Biophys. J. 101 (10) (2011) 2525–2534.
[42] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, J. Appl. Crystallogr. 26 (1993) 283–291.
[43] M. Jamroz, A. Kolinski, S. Kmiecik, CABS-flex: server for fast simulation of protein structure fluctuations, Nucleic Acids Res. 41 (W1) (2013) W427–W431.
[44] R.R. Burgess, Chapter 17 Refolding Solubilized Inclusion Body Proteins. Guide to Protein Purification, second ed., (2009), pp. 259–282.
[45] A. Basu, X. Li, S.S.J. Leong, Refolding of proteins from inclusion bodies: rational design and recipes, Appl. Microbiol. Biotechnol. 92 (2) (2011) 241–251.
[46] A. Ramón, M. Señorale-Pose, M. Marín, Inclusion bodies: not that bad…, Front. Microbiol. 5 (2014) 56.
[47] J.N. Andersen, P.G. Sriram, N. Kalogerakis, L.A. Behie, Effect of temperature on recombinant protein production using the Bm5/Bm5.NPV expression system, Can. J. Chem. Eng. 74 (4) (1996) 511–517.
[48] A. Farewell, F.C. Neidhardt, Effect of temperature on in vivo protein synthetic capacity in Escherichia coli, J. Bacteriol. 180 (17) (1998) 4704–4710.
[49] H.P. Sørensen, K.K. Mortensen, Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli, Microb. Cell Factories 4 (2005) 1.
[50] Y. Tsunoda, N. Sakai, K. Kikuchi, S. Katoh, K. Akagi, J. Miura-Ohnuma, ... E. Katoh, Improving expression and solubility of rice proteins produced as fusion proteins in Escherichia coli, Protein Expr. Purif. 42 (2) (2005) 268–277.
[51] N. Akbari, K. Khajeh, N. Ghaemi, Z. Salemi, Efficient refolding of recombinant lipase from Escherichia coli inclusion bodies by response surface methodology, Protein Expr. Purif. 70 (2) (2010) 254–259.
[52] M.J. Weickert, M. Pagratis, S.R. Curry, R. Blackmore, Stabilization of apoglobin by low temperature increases yield of soluble recombinant hemoglobin in Escherichia coli, Appl. Environ. Microbiol. 63 (11) (1997) 4313–4320.

[53] A.L. Larentis, J.F.M.Q. Nicolau, G. dos S. Esteves, D.T. Vareschini, F.V.R. de Almeida, M.G. dos Reis, ... M.A. Medeiros, 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 micro bioreactor, BMC Res. Notes 7 (2014) 671.

[54] G. Rabbani, J. Kaur, E. Ahmad, R.H. Khan, S.K. Jain, Structural characteristics of thermostable immunogenic outer membrane protein from Salmonella enterica serovar Typhi, Appl. Microbiol. Biotechnol. 98 (6) (2013) 2533–2543.

[55] G. Rabbani, E. Ahmad, M.V. Khan, M.T. Ashraf, R. Bhat, R.H. Khan, Impact of structural stability of cold adapted Candida Antarctica lipase B (CaLB): in relation to pH, chemical and thermal denaturation, RSC Adv. 5 (26) (2015) 20115–20131.

[56] G. Rabbani, E. Ahmad, N. Zaidi, R.H. Khan, pH-dependent conformational transitions in conalbumin (ovotransferrin), a metalloproteinase from hen egg white, Cell Biochem. Biophys. 61 (3) (2011) 551–560.

[57] A. Varshney, B. Ahmad, G. Rabbani, V. Kumar, S. Yadav, R.H. Khan, Acid-induced unfolding of didecameric keyhole limpet hemocyanin: detection and characterizations of decameric and tetrameric intermediate states, Amino Acids 39 (3) (2010) 899–910.