



Optimization of induction parameters, structure quality assessment by ATR-FTIR and *in silico* characterization of expressed recombinant polcalcin in three different strains of *Escherichia coli*

Afshin Derakhshani ^{a,b}, Kamyar Keshavarz F. ^{c,d}, Sahar Barzegari Banadkoki ^c, Farshad H. Shirazi ^{c,d}, Mahmood Barati ^e, Mohammad Fereidouni ^{f,*}, Hossein Safarpour ^{f,*}

^a Student Research Committee, Birjand University of Medical Sciences, Birjand, Iran

^b Immunology Research Center, Tabriz University of Medical Sciences, Tabriz, Iran

^c Pharmaceutical Sciences Research Center, Shahid Beheshti University of Medical Sciences, Niayesh Highway, Valiasr Ave, Tehran, Iran

^d Department of Toxicology/Pharmacology, School of Pharmacy, Shahid Beheshti University of Medical Sciences, Niayesh Highway, Valiasr Ave, Tehran, Iran

^e School of Allied Medical Sciences, Iran University of Medical Sciences, Tehran, Iran

^f Cellular and Molecular Research Center, Birjand University of Medical Sciences, Birjand, Iran

ARTICLE INFO

Article history:

Received 23 May 2019

Received in revised form 10 July 2019

Accepted 10 July 2019

Available online 11 July 2019

Keywords:

Recombinant allergen

Polcalcin (Che a 3)

Response Surface Methodology

ATR-FTIR structural analysis

In silico characterization

ABSTRACT

The inhalation of *Chenopodium album* (*C. album*) pollen, especially polcalcin (Che a 3), has been reported as a significant source of allergic respiratory symptoms. This study was conducted to characterize biophysical differences of recombinant polcalcin come from three different *Escherichia coli* strains using attenuated total reflectance Fourier transform infrared spectroscopy (ATR-FTIR), optimize recombinant polcalcin expression, and linear B-cell epitopes identification using *in silico* methods. ATR-FTIR results of purified proteins showed spectra intensity variance in the amide I region, while there were no changes in peak position and shape of the bands. SHuffle® T7 Express was selected for subsequent optimization due to ability in the correction of the mis-oxidized bonds and promotes proper folding which was validated by ATR-FTIR analysis results. Then, Response Surface Methodology was performed to optimize critical factors including induction temperature, duration of induction, and concentration of inducer. The best partitioning conditions were achieved in 1.5 mM IPTG for 10.97 h at 29.9 °C. Finally, prediction of polcalcin B-cell epitopes was carried out which indicated the presence of 4 different epitopes. Together, the results may help to the development of diagnostic approaches and also vaccine manufacture for desensitization and modulation of the allergic response in patients.

© 2019 Published by Elsevier B.V.

1. Introduction

During the recent decades, there has been a huge increase in the incidence of allergic diseases [1]. Allergic diseases which are a group of immune-mediated disorders, generally triggered by an IgE-dependent immunological response to an allergen [2–4]. Although the pattern of allergic sensitization is varied in different areas, plant pollens especially pollen of weeds and grasses are the most common triggers of allergic symptoms [5]. Regarding the high prevalence of allergic disorders around the world and its chronic nature, allergies can negatively affect patient's quality of life and imposed a huge economic burden on societies and health systems [6,7], therefore proper diagnosis and management of allergies is a very important issue.

Currently, the most common diagnostic methods for allergic disorders are skin prick test and detection of allergen-specific IgE in blood [8]. In both tests, natural allergen extracts containing major and minor epitopes of each allergen such as pollens, foods, molds and animal's dander are used. In all *in vivo* and *in vitro* diagnostic tests for allergy, the source of allergenic materials, quality, and quantity of extracts are crucial and affect the sensitivity and specificity of the test [9]. However, studies have shown that there is a large variation in the quality of commercial allergen extracts [10–12]. Rather than the difference in the source materials, presence of allergenic and non-allergenic ingredients, lack of standardization and variation in extraction methods decrease the uniformity and validity of natural extracts [10–13]. Furthermore, many inhalant and food allergens have intrinsic enzyme activities which make them susceptible to degradation in a time-dependent manner and causes a reduction in the allergenicity as well as the biological activity of the extracts [9]. Besides the problems mentioned above, natural allergenic extracts are expensive and are prone to batch-to-batch variation and also cannot be constructed to suit the individual patient's

* Corresponding authors.

E-mail addresses: dr.m.fereidouni@gmail.com (M. Fereidouni), H.safarpour@bums.ac.ir (H. Safarpour).

sensitization profile. An alternative way to overcome the disadvantages and problems associated with natural allergenic extracts is to use recombinant allergens [14]. During the last three decades, utilization of recombinant DNA technology in the field of allergy diagnosis and treatment has led to the production of pure, reproducible, consistent and physicochemically defined allergens with low cross-reactivity at a better cost.

There are many hosts used for the production of recombinant protein such as bacteria, yeast, plants, and animals [15,16]. *Escherichia coli* (*E. coli*), is one of the most important expression systems which is used routinely for rapid production of a large amount and low-cost recombinant allergens production [14,17]. Notably, recombinant protein expression in *E. coli* often results in the formation of insoluble nonfunctional inclusion bodies [18]. In general, many factors such as post-induction temperature [19], cultivation media, *E. coli* strain, and inducer concentration [20] affecting expression and solubility rates which can be manipulated in order to obtain an optimum condition for overexpression and reduce inclusion bodies formation.

Pollens are the most common triggers of allergic symptoms around the world. Among different type of allergenic pollens, *Chenopodiaceae* family and particularly *C. album* is one of the most important causes of allergic sensitization. Several studies from different parts of Iran [21,22] and other countries [23–25] confirmed the high rate sensitization to *C. album* allergens among allergic patients.

Based on our previous study which revealed the role of expression system type on secondary structure differences [26], we first aimed to characterize physicochemical properties of recombinant polcalcin expressed in three types of *E. coli* strains including SHuffle® T7 Express, BL21 (DE3) and Rosetta-gami™ 2 (DE3) through ATR-FTIR. In the next step, we compared the effects of different cultivation conditions using Response Surface Methodology (RSM) to reach the maximum yield of recombinant polcalcin. In this regard, three parameters including induction periods and temperature as well as inducer concentration were examined.

As IgE binding sites characterization can be used for the discovery and development of new therapeutics, vaccines, and diagnostics of many allergic diseases, using updated algorithms, an epitope mapping process was finally conducted to determine IgE binding sites of polcalcin.

2. Material and methods

2.1. Strains, vector and culture media

The recombinant pET21b (+) vector containing polcalcin encoding gene was kindly provided by Dr. Sankian [27]. Production of high-efficiency chemical competent cells was conducted using the standard protocols [26]. The heat shock method was used for transformation of the bacterial cells and the transformed clones were selected using LB agar medium containing 100 µg mL⁻¹ of ampicillin. Aliquots containing transformed bacterium were stored in 20% (v/v) glycerol at -70 °C for long term usage.

2.2. Expression of polcalcin in different strains of *E. coli*

A loopful of bacteria were inoculated into 10 mL of LB broth medium and grown overnight in shaking incubator at 37 °C. A total volume of 1 mL of the overnight culture was transferred to the falcons containing 100 mL fresh sterilized LB medium and 100 µg mL⁻¹ ampicillin. Cells were grown at 37 °C and 200 rpm till they reached OD_{600nm} of 0.4–0.6. The recombinant protein expression was then induced by the addition of IPTG (Isopropyl-β-D-thiogalactopyranoside) to the final concentrations of 1 mM. The cultures were incubated further at 28 °C for 4 h. After that, the cells were harvested by centrifugation at 7000g for 20 min at 4 °C. Finally, the pellets of induced bacteria were analyzed

by 12% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE).

2.3. Polcalcin purification

The clarified supernatant containing soluble His-tagged recombinant polcalcin protein was added to the immobilized metal affinity chromatography (IMAC) column with Ni-NTA agarose beads (Qiagen, Netherlands). The gravity-flow-based chromatography was conducted under native conditions (Qiagen, Netherlands). Finally, the purified recombinant polcalcins from three different strain were quantified by the Bradford method [28] and were analyzed by SDS-PAGE for further ATR-FTIR characterization.

2.4. ATR-FTIR analysis

2.4.1. Data acquisition

FTIR data acquisition was done by two methods using WQF-510 Fourier transform spectrometer (Rayleigh Optics, China) equipped with a DLATGS (deuterated, L-alanine doped triglycine sulfate) detector. First, the background of ZnSe window cell (13 mm × 3 mm) was obtained. Then, the spectrum of elution buffer and protein in aqueous phase were acquired separately. Second, to confirm the results of the first method, concentration-corrected volumes of recombinant polcalcin and buffer were undergone using harsh freeze-drying condition. Having further acquired the background spectrum of the diamond crystal and air, the IR spectra of identical volumes of freeze-dried powders comprising buffer and recombinant polcalcin in different hosts, was obtained using MIRacle™ single reflection attenuated total reflection (ATR) (PIKE technologies, WQF-510, Rayleigh optics, China) with window made of diamond crystal. Raw spectra were acquired in the 400 to 4000 cm⁻¹ spectral range with a resolution of 4 cm⁻¹. To avoid low signal to noise ratio (SNR), 512 scans were co-added for each spectrum. Experiments were done triplicate. Spectra were collected against the background prior to each measurement. The chamber of spectrophotometer was inflated with N₂ gas during data acquisition.

2.4.2. Spectral preprocessing and data analysis

Essential preprocessing and statistically analysis of raw FTIR spectra were performed by MATLAB software version 2018b (The Math-Works, Natick, USA). First and foremost, the spectrum of buffer was subtracted from the spectra of polcalcin. Second, Amide I region (1600–1700 cm⁻¹) were cut. Then, to smooth the peaks, Savitzky_Golay filter with a fourth order polynomial in fifteen points and averaging filter function were applied to the mean of triplicate spectra. Fourth, min-max normalization (0–1) was performed. For quantitative analysis of the secondary structure of polcalcin, the areas under the peaks of deconvolved spectra were calculated with respect to peak assignments mentioned in [29] and divided by the sum of all peak areas to obtain the relative percentage of each peak. In addition, the second derivative of deconvolved spectra was used to clearly identify and compare each peaks dimensions to investigate peak generation, peak removal or shifts.

2.5. Experimental design and optimization by RSM

RSM as an empirical statistical modeling technique was employed to optimize the expression conditions that enhanced the recombinant polcalcin production. Central composite design (CCD) developed by the Design Expert software version 6.0.8 (Stat-Ease, Minneapolis, USA), was used to optimize the three significant cultivation conditions in SHuffle® T7 Express: post-induction temperature, inducer concentration and post-induction period. A set of 20 experimental runs with 5 replicated center points were generated. Three different levels: low, medium and high were used to study the independent variables (Table 1). The experimental design used for the study is presented in

Table 1
Independent variables and their coded and actual values used for the optimization.

Factor	Name	Units	Type	Low	Mean	High
A	IPTG	mM	Numeric	0.5	1	1.5
B	Induction time	h	Numeric	4	8	12
C	Temperature	°C	Numeric	16	28	37

Table 2. All the experiments were carried out in triplicates and the average of polcalcin yield was considered as the response. Protein quantification was performed by scanning-based densitometry analysis of acrylamide gels using ImageJ software (NIH, MD, USA).

2.6. IgE immunoassay with dot blot

Ten micrograms of purified recombinant polcalcin protein, as well as Bovine Serum Albumin (BSA) and natural extract of *C. album* (Holishter-steer, USA) as the negative and positive controls, respectively were blotted onto nitrocellulose membrane allowed to dry for 1 h at room temperature. After blocking of unreacted protein-binding sites by skimmed milk (2%), the membrane was washed three times (3 min per wash) with TBST buffer (20 mM Tris and 150 mM NaCl, pH 7.6, 0.1% Tween-20). Next, membranes were incubated with a pooled sera of five allergic patients which showed positive skin prick test to the commercial crude extract of *C. album* (Holishter steer, USA) for 1 h at room temperature and washed 3 times with TBST buffer. Then, enzyme-labeled anti-IgE antibody (KPL, the Netherlands) was added to the membrane and incubated at 37 °C for 1 h, followed by washing with washing buffer and a final round of washing with ddH₂O. Freshly prepared substrate solution of tetramethyl benzidine (TMB) (Sigma,

Switzerland) was added onto the membrane for color development for 10–30 min.

2.7. In silico characterization of polcalcin protein

The prediction of B-cell epitopes was performed to find potential epitopes that can cause an allergic reaction in the susceptible people. The three-dimensional crystal structure of polcalcin (PDB ID: 2OPO) was retrieved from RCSB PDB database and linear B-cell epitope prediction was conducted using BepiPred modified hidden Markov model from IEDB analysis resource v2.12 (<http://tools.immuneepitope.org/bcell/>) [30]. The recommended cutoff of 0.35 was used to determine potential B-cell linear epitopes, ensuring sensibility of 49% and specificity of 75% to this approach. Linear B-cell epitopes are predicted to be located at the residues with the highest scores. The Emini surface accessibility (ESA) was used to evaluate the probability of predicted linear B-cell epitopes to be exposed on the surface of the protein. This approach calculates the surface accessibility of hexapeptides and values >1.0 indicate an increased probability of being found on the surface (36). Sequences with BepiPred score above 0.35 and ESA score above 1.0 were considered potential linear B-cell epitopes in regions that could be accessed by naturally acquired antibodies.

3. Result

3.1. Protein expression and molecular weight determination

SHuffle® T7 Express, BL21 (DE3) and Rosetta-gami™ 2 (DE3) cells were transformed with recombinant pET21b-polcalcin plasmid and induced with 1 mM of IPTG at 28 °C. The SDS-PAGE analysis after induction confirmed the presence of recombinant protein with a major

Table 2
Central composite design (CCD) matrix with the values of the independent variables IPTG concentration (factor A), induction time (factor B), induction temperature (factor C), and the results for the production of polcalcin.

	Factor A	Factor B	Factor C	Response 1	Response 2
Run	IPTG (mM)	Induction Time (h)	Temperature (°C)	Protein (µg/µL)	Cell Density (OD ₆₀₀)
1	1	8	27.5	745	1.56
2	1.5	4	37	334	1.185
3	1	8	27.5	740	1.62
4	1	1.27	27.5	50	0.499
5	1	8	27.5	710	1.62
6	1.5	12	37	1151	2.52
7	1	8	27.5	701	1.62
8	0.5	12	18	848	0.906
9	0.5	4	18	70	0.6565
10	1.5	12	18	1250	0.828
11	1.5	4	18	115	0.628
12	1	8	27.5	860	1.74
13	1	8	27.5	803	1.695
14	1	14.73	27.5	1350	2.705
15	0.5	12	37	1034	2.605
16	1.84	8	27.5	1364	1.68
17	0.16	8	27.5	86	1.755
18	1	8	11.52	283	0.16
19	1	8	43.48	365	1.845
20	0.5	4	37	346	1.225

The gradient color is related to protein concentration from white (minimum=86 µg/µL) to red (maximum=1364 µg/µL).

band at about 14 kDa, which was consistent with the expected molecular weight of the polcalcin (Fig. 1A). The quantity of polcalcin produced by the strains were quantified by densitometry of the corresponding bands of stained SDS-PAGE gels. SHuffle® T7 Express and BL21 (DE3) produced significantly more proteins than the Rosetta-gami™ 2 (DE3) cells. On the other hand, no significant differences in protein expression were observed for the SHuffle® T7 Express and BL21 (DE3) cells.

3.2. Recombinant polcalcin purification

Purification was performed using Ni-NTA columns in IMAC method and analyzed by SDS-PAGE. As shown in Fig. 1B, successful expression and purification of recombinant polcalcin with a molecular weight of about 14 kDa was observed. The result of Bradford method has also revealed that the concentration of polcalcin were about 500, 575 and 600 $\mu\text{g mL}^{-1}$ in *E. coli* Rosetta-gami™ 2 (DE3), SHuffle® T7 Express and BL21 (DE3) cells, respectively.

3.3. ATR-FTIR characteristics

The primary goal of this study was to test the feasibility of FTIR spectroscopy to acquire information regarding the differences of the secondary structure of recombinant polcalcin which expressed in SHuffle® T7 Express, BL21 (DE3) and Rosetta-gami™ 2 (DE3) cells. FTIR as a rapid, non-destructive and easy to perform analytical method gives high-quality spectra using a very small amount of protein. It is especially useful for the analysis of the secondary structural characterization of proteins because characteristic protein bands (e.g. the amide I region) in an IR spectrum are very sensitive to conformational changes in the protein [31].

In order to the quantitative determination of secondary structural changes, the curve-fitting analysis was applied to the amide I region ($1700\text{--}1600\text{ cm}^{-1}$). The contribution of each component band to the amide I band is shown in Figs. 2 and 3. Second derivatives of deconvoluted SHuffle® T7 Express, BL21 (DE3) and Rosetta-gami™ 2 (DE3) cell spectra were calculated for the evaluation of shift position, peaks generations, and their removal. The FTIR results suggested negligible secondary conformational modifications among the various proteins expressed in three types of *E. coli*. As shown in Fig. 2, these structural differences were limited only to spectral intensity variance in the amide I. Furthermore, pick position and shape of the bands were not affected by the type of *E. coli* strain. Also, to verify these

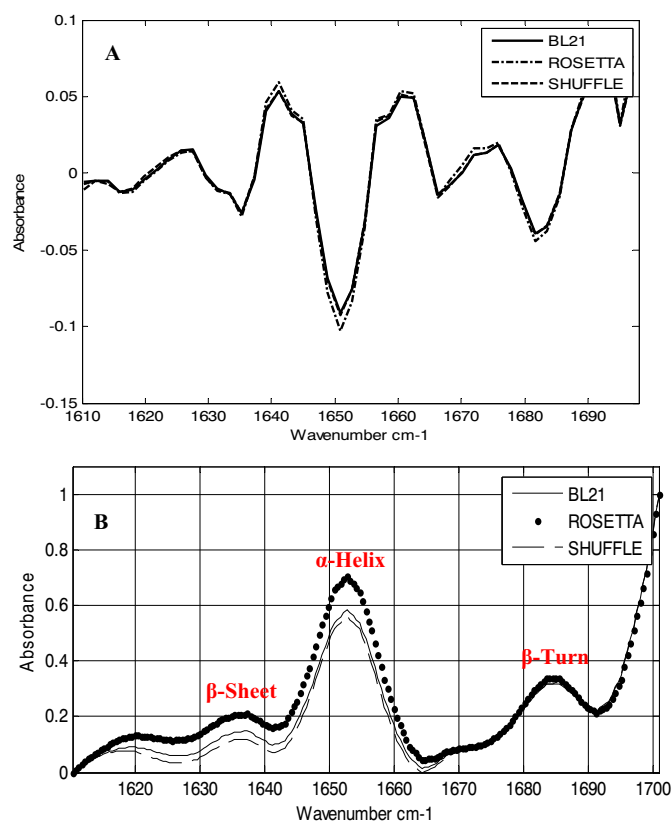


Fig. 2. FTIR characterization of liquid polcalcin. A: Secondary derivative of SHuffle® T7 Express, Rosetta-gami™ 2 (DE3) and BL21 (DE3) spectra; B: display variance of absorption intensities.

discriminating differences and similarities, a freeze-drying process was conducted on proteins. The result of ATR indicated no significant differences were observed between the spectra of three strain (Fig. 3). Also, the results of the comparison between ATR and FTIR indicated the presence of a peak in ATR at 1670 cm^{-1} which was assigned to β -Sheet (Fig. 3). Moreover, a peak which observed at 1620 cm^{-1} which can be attributed to nucleic acid due to the base carbonyl stretching and ring breathing mode (Table 4) [32]. The peak in this position is

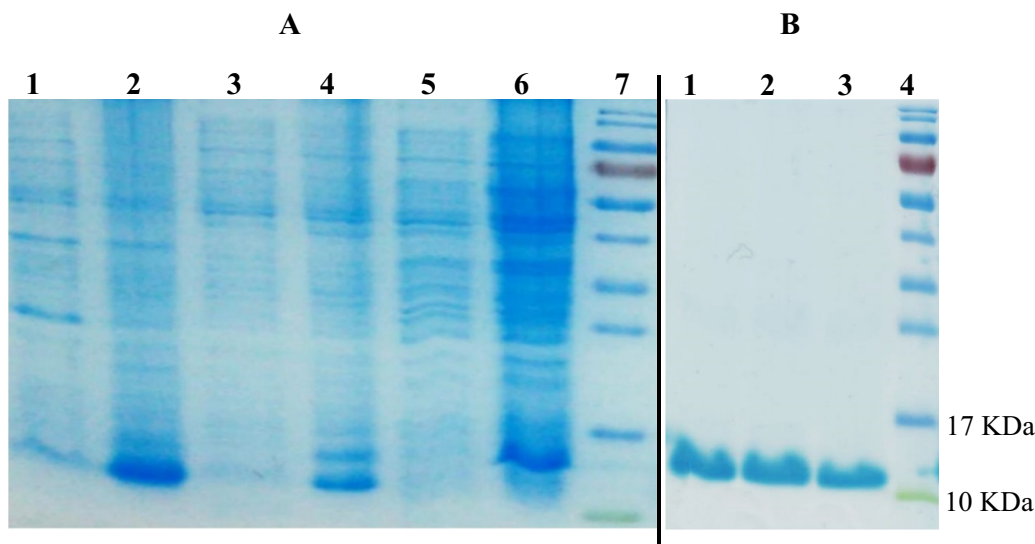


Fig. 1. Analysis of polcalcin expression in three different *E. coli* strains using SDS-PAGE 12.5%. A: Lanes 1 and 2: bacterial protein before and after induction with IPTG respectively (BL21 (DE3)); Lanes 3 and 4: bacterial protein before and after induction with IPTG respectively (SHuffle® T7 Express); Lanes 5 and 6: bacterial protein before and after IPTG induction respectively (Rosetta-gami™ 2 (DE3)); Lane 7: Pre-stained Protein Ladder (Thermo Fisher Scientific).

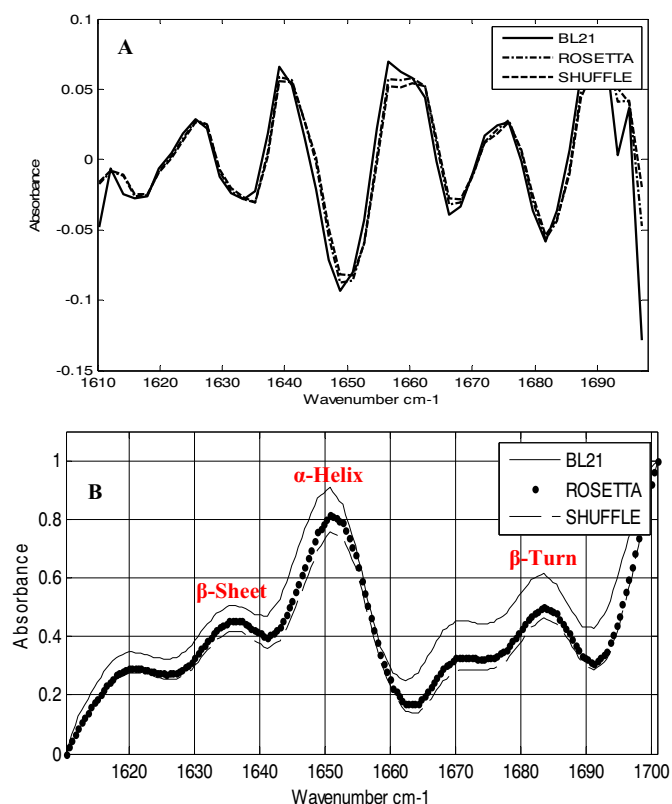


Fig. 3. ATR characterization of lyophilized polcalcin. A: Secondary derivative of SHuffle® T7 Express, Rosetta-gami™ 2 (DE3) and BL21 (DE3) spectra; B: display variance of absorption intensities.

rarely seen for proteins that may be related to contamination with nucleic acids. Quantitative ATR-FTIR analysis of the amide I band shows the secondary structure of polcalcin has some minimal differences between the SHuffle® T7 Express and the Rosetta-gami™ 2 (DE3) compared to BL21 (DE3) (Table 5).

3.4. Optimization of polcalcin production

Determination of optimum concentrations of IPTG (A), induction time (B) and induction temperature (C) to enhance the cytoplasmic polcalcin production in SHuffle® T7 Express was carried out using the CCD design. The yield of protein responses exhibited a range of 50–1364 $\mu\text{g mL}^{-1}$ in different conditions. The experimental design with the actual levels of the variables and the results of experiments are shown in Table 2. The observed results were fitted with a second-order polynomial function for estimation of polcalcin production (protein) in terms of coded factors as follows:

$$\text{Protein} = 564.05333 - 1970.71004A - 543.29313B + 90.45872C + 652.27835AB + 38.46938B^2 - 1.54382C^2 - 38.86896AB^2$$

where A, B, and C are IPTG concentration, induction time and induction temperature, respectively.

Regression analysis of the experimental design for linear, quadratic, 2FI and cubic was performed. The R^2 value of 0.97498 and the adjusted R^2 value of 0.960386 indicated a good fitness of the model which can explain 91.2% of the variability in the response.

The relationship between the observed polcalcin production and the expected values by the model is also represented in Fig. 4. The results of ANOVA test (Table 3) with the model F value of 66.8032 and a very low probability value also indicated the statistical significance and adequacy of the cubic model. As shown in Table 3, all single factor terms (A, B and

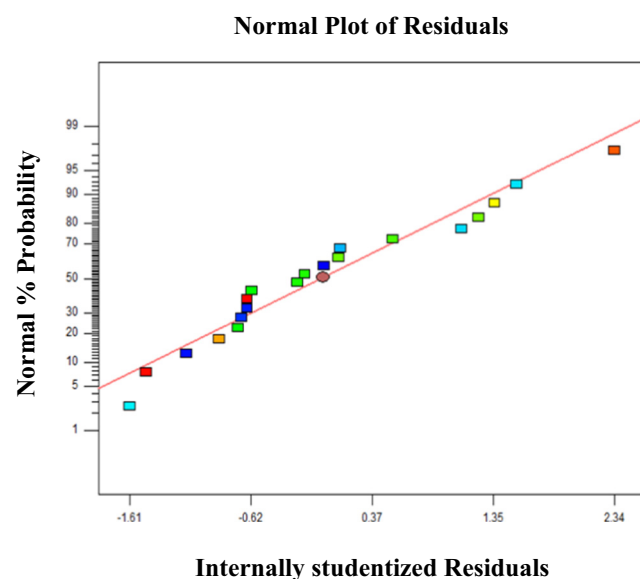


Fig. 4. Predictability of the obtained model for recombinant polcalcin overexpression was studied using the scatter plot of the predicted versus measured experimental densitometric values.

C) and AB were the main significant variables affecting polcalcin production (p -value < 0.05) while other interaction and cubic model terms had no significant effect (p -value > 0.05) on polcalcin production.

To visualize the effects of the variables on the response and evaluate the interactions and optimum levels of the variables for polcalcin production, the response surface plot was generated for each pair of experimental factors, while the other two were held constant at their middle levels. Fig. 5 illustrates the surface plots of polcalcin production for the interactive effect of all possible pairwise combinations of variables. The analysis of the plots demonstrated that the highest polcalcin yield was achieved when the IPTG concentration and incubation time were adjusted to 1.5 mM and 12 h, respectively. It was also observed that the increase in IPTG concentration up to 1.5 mM enhanced the polcalcin yield but higher concentrations of IPTG had no effect on the expression yield. The highest polcalcin yield predicted by the model was 1319 $\mu\text{g mL}^{-1}$ with optimum values of IPTG concentration (1.5 mM), induction time (10.97 h), temperature (29.3 °C). Consequently, the validation experiments were carried out under the optimized condition and the yield of 1226 $\mu\text{g mL}^{-1}$ was achieved which was very close to the value predicted by the model. This yield was obtained in the optimized medium with constant induction condition of 1.5 mM IPTG and 10.97 h post-induction at 29.3 °C. The result of SDS-PAGE analysis showed in

Table 3
ANOVA for response surface reduced cubic model.

Source	df	Sum of squares	Mean square	F Value	p-Value	
Model	7	3,505,150	500,735.6	66.8032	<0.0001	Significant
A-IPTG	1	816,642	816,642	108.9483	<0.0001	
B-induction time	1	2,299,836	2,299,836	306.8213	<0.0001	
C-temperature	1	37,949.16	37,949.16	5.062801	0.0440	
AB	1	29,524.5	29,524.5	3.938867	0.0705	
B ²	1	594.9199	594.9199	0.079368	0.7830	
C ²	1	282,547.5	282,547.5	37.69469	<0.0001	
AB ²	1	320,405.6	320,405.6	42.74534	<0.0001	
Residual	12	89,948.21	7495.684			
Lack of fit	7	71,493.38	10,213.34	2.767118	0.1401	Not significant
Pure error	5	18,454.83	3690.967			
Cor total	19	3,595,098				

Table 4
Intensity discrimination of variance of absorption.

Lyophilized polcalcin		Liquid polcalcin	
Peak cm^{-1}	Assignment [32]	Peak cm^{-1}	Assignment [32]
1600–1620	The region of the amide I band β -Sheet structure of amide I Proportion of b-sheet secondary structures (shoulder)	1650–1663	Amide I (α -helix)
1625–1647			
1660–1668	Peak of nucleic acids due to the base carbonyl stretching and ring breathing mode	1640–1649	Unordered random coils and turns of amide I Amide I (anti-parallel β -sheet)
1621–1624			
1669–1700	Amide I (anti-parallel β -sheet)	1664–1695	Amide I (disordered structure-solvated) Amide I (anti-parallel β -sheet) Antiparallel β -sheet of amide I
1648–1659	Unordered random coils and turns of amide I	1600–1639	β -Sheet structure of amide I Proportion of b-sheet secondary structures (shoulder) Amide I region

Fig. 6 and clearly indicated that a higher amount of protein after optimization.

3.5. Reactivity of recombinant polcalcin with patients IgE

IgE reactivity of recombinant polcalcin which produced by SHuffle® T7 was checked by an immunoblotting technique using pooled sera of sensitized patients. As shown in Fig. 7, both recombinant and native forms of polcalcin, showed a positive result while no IgE binding was detected for negative control (BSA). The results indicated that SHuffle® T7 Express can produce a bio-active form of recombinant polcalcin.

3.6. Mapping of IgE-binding epitopes

Knowledge of the allergen epitope characteristics is pivotal for understanding the pathogenesis of an allergen. Here, IgE binding regions were identified by sequence-based tools which fall in common regions of the sequence map. Three stretches of amino acids 1–12, 21–34, 42–48 and 55–62 recognized by two tools were predicted as B-cell epitopes. The result of epitope mapping is presented in Fig. 8.

4. Discussion

Recombinant allergens have considerably advanced our knowledge of the immune mechanisms of allergic diseases as well as vastly enriched our toolbox for diagnosis and therapy. In this study, we have successfully established an optimized-condition for the production of recombinant polcalcin allergen in *E. coli* that is pivotal for developing detection, diagnostic, and treatment methods.

In order to express a recombinant protein in high yield soluble form, several strategies like the introduction of molecular chaperons to the host bacteria, modulation of expression level, optimization of culture medium, reduction of temperature, the addition of chemical chaperons have been introduced [33]. With thanks to advances in genetic engineering, recently some new *E. coli* strains such as SHuffle® T7 Express and Rosetta-gami™ 2 (DE3) have been developed which capable of

expressing recombinant proteins with intact structure. Previous studies showed that these *E. coli* strains can express a particular protein in different structures compared to older strains e.g. BL21 (DE3) [26,34]. Hence, in order to find the best strain of *E. coli* for soluble expression and high yield production of polcalcin, the encoding gene was cloned and successfully expressed in these three different strains.

We revealed that expression yield of soluble polcalcin in SHuffle® T7 Express was almost equal to BL21 (DE3) but much higher than Rosetta-gami™ 2 (DE3). The significant increase in expression of soluble polcalcin in SHuffle® T7 Express in compare with Rosetta-gami™ 2 (DE3) was interesting, as the protein has no disulfide bond. It suggests that the increased expression may not be related to the SHuffle® T7 Express cytoplasm's oxidizing environment, as the weak expression was demonstrated in Rosetta-gami™ 2 (DE3) which has an oxidizing cytoplasmic environment. Limited studies described the expression of non-disulfide bonded proteins in SHuffle® T7 Express [35,36]. For example, Ren et al., showed increased expression of soluble GFP with two free cysteines, in SHuffle® strain [35], while expression in other strains leads to the formation of insoluble GFP protein [35]. It is hypothesized that cytoplasmic disulfide bond isomerase C (DsbC), in SHuffle® T7 Express, enhances correct folding of GFP by its chaperone property but not disulfide isomerase activity. So, DsbC may have a contribution in protein folding via an activity which is distinct from its redox activity [37].

The produced recombinant allergens were then purified and used for ATR-FTIR characterization. As polcalcin is a small protein with 86 amino acids which has no disulfide bond or other complex domains, ATR-FTIR characterization results have indicated minor structural and conformational changes between the different proteins expressed in *E. coli* strains. These structural differences were limited to spectral intensity variance, while there were no changes in pick position and shape of the bands. Most of these discriminating differences and similarities are characterized by amide I bands. To further confirm these results, proteins were gone through the harsh conditions of freeze-drying. However, no significant differences have been seen between spectra which might have been raised from structural instabilities in these conditions. Quantitative analysis of the amide I band shows the secondary structure of polcalcin has changed between the SHuffle® T7 Express and the Rosetta-gami™ 2 (DE3) compared to BL21 (DE3). Similar observations were previously reported in other proteins containing β -sheet. For example, we have previously shown significant differences between β -sheet secondary structures of recombinant Tumor Necrosis Factor α in two *E. coli* strains [26]. Whether these minimal differences in structure might have significant effects on polcalcin allergenicity need further investigations. Thus, the results of FTIR suggest that there is no priority to use which strain for the production of polcalcin.

In the next step, we used RSM for the optimization of high yield recombinant polcalcin production. The RSM has been proved to be an effective tool for identifying significant variables and experimental

Table 5
Percentage distributions of secondary structures of polcalcin expressed in three different strains determined by ATR-FTIR spectroscopy.

Sample	<i>E. coli</i> strain	Secondary structure (overall %)		
		β -Turn	α -Helix	β -Sheet
Liquid polcalcin	SHuffle® T7 Express	31.80	35.20	34.01
	BL21 (DE3)	28.37	37.89	33.92
	Rosetta-gami™ 2 (DE3)	27.09	38.51	35.06
Lyophilized polcalcin	SHuffle® T7 Express	31.04	39.70	30.57
	BL21 (DE3)	32.46	38.18	30.44
	Rosetta-gami™ 2 (DE3)	31.47	39.56	30.56

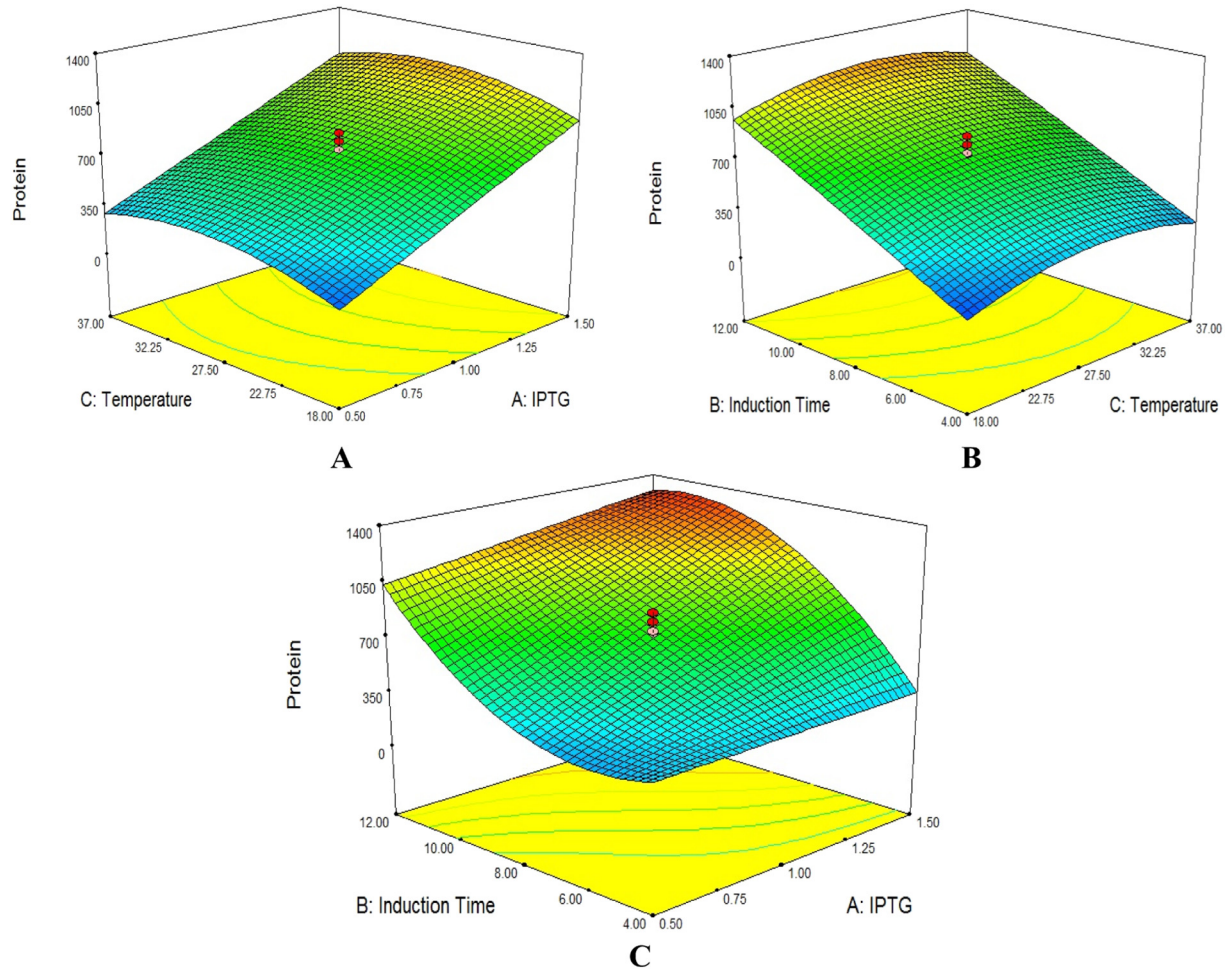


Fig. 5. Response surface plots (3D) showing the influence of three studied variables on the polyclonal production (A) interaction between IPTG concentration and induction temperature; (B) interaction between induction time and induction temperature; and (C) interaction between IPTG concentration and induction time.

conditions to optimize the expression of several proteins by varying different factors [38]. In our study, the highest increase of protein production in SHuffle® T7 Express was achieved after the determination of the optimum induction conditions of 1.5 mM IPTG and 10.97 h post-induction at 29.3 °C. The obtained amount of polyclonal production under optimized conditions showed 92.9% productivity, showing a good coincidence with the predicted value, and thus indicating the adequacy of the fitted model.

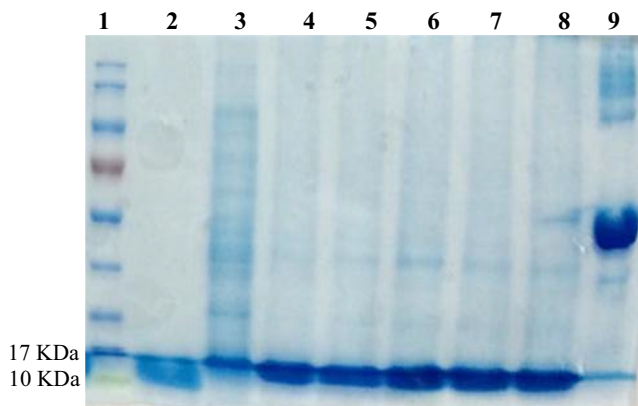


Fig. 6. SDS-PAGE analysis of optimized condition samples. Lane 1: Pre-stained Protein Ladder (Thermo Fisher Scientific); lane 2: purified polyclonal; lane 3: non-optimized polyclonal; lane 4–8: condition optimized polyclonals; and lane 9: BSA (1 mg mL⁻¹).

IgE binding of native and recombinant forms of polyclonal was recognized with sera from allergic individuals by immune-blotting, providing evidence of the protein as an allergen. In order to find which epitopes are involved in IgE reactivity of polyclonal, an epitope mapping process was conducted using novel algorithms. More evidence regarding the amino acid level of epitopes recognized by IgE is important for their clinical application and also for understanding general molecular recognition events [39]. Generally, the peptide that has at least 15 to 20 mer is a good candidate for the development of specific enzyme immunoassays (EIA) based immunodiagnostic tools [40]. Instead of overlapping synthetic peptides used to validate IgE reactivity in previous studies, epitope mapping by immune-informatics tools may decrease the number of synthesized peptides and working time [41]. However, in the case of polyclonal as a major allergen, to the best of our knowledge, no study has not been done until now. The B-cell epitope identified in this study has potential application for the detection of IgE antibodies in

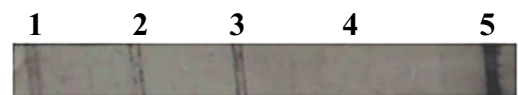


Fig. 7. Dot blot assay. 1, 2, and 3: Purified recombinant polyclonal expressed in SHuffle® T7 Express, Rosetta-gami™ 2 (DE3) and BL21 (DE3) respectively; 4: Negative control (BSA); and 5: Native polyclonal; B: Lane 1: Pre-stained Protein Ladder (Thermo Fisher Scientific). Lanes 2, 3, and 4: purified recombinant polyclonal expressed in SHuffle® T7 Express, Rosetta-gami™ 2 (DE3) and BL21 (DE3) respectively.

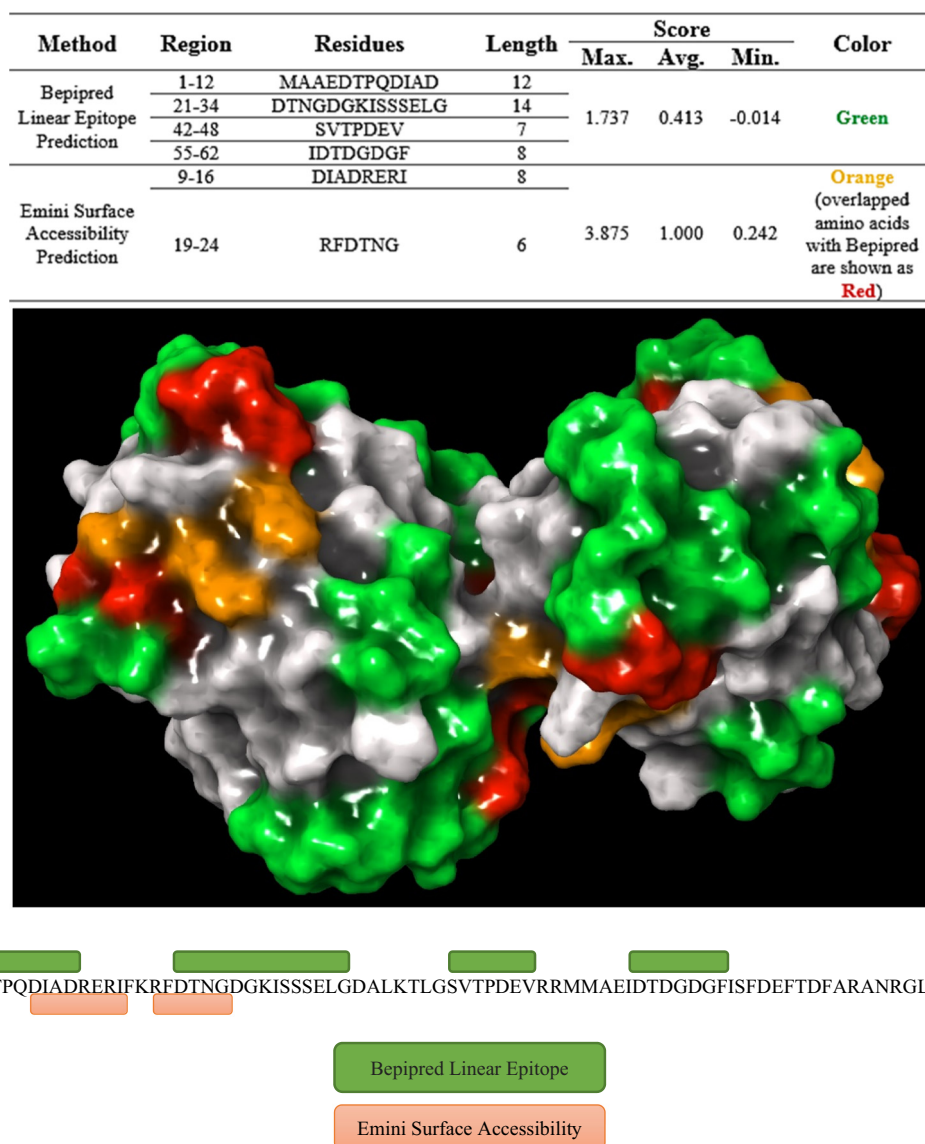


Fig. 8. Predicted epitope residues of the polcalcin.

ELISA or point of care test for diagnosis. However, this needs to be evaluated using a panel of sera.

5. Conclusion

The present study described the high yield expression and physico-chemical characterization of polcalcin as well as *in silico* identification of epitopes. Our experimental results not only provide a basis for large scale production of recombinant polcalcin for clinical diagnostic purposes but also with the aim of predicted epitopes have significance in experimental research and immunotherapy.

Declaration of Competing Interest

The authors declare that there is no conflict of interest.

References

- [1] W. Eder, M.J. Ege, E. von Mutius, The asthma epidemic, *N. Engl. J. Med.* 355 (21) (2006) 2226–2235.
- [2] P. Ilaria, L.-S. Francesca, M. Paola, Chapter 5 - therapeutic approaches in allergic diseases A2 - Boraschi, Diana, in: G. Penton-Rol (Ed.), *Immune Rebalancing*, Academic Press 2016, pp. 85–103.
- [3] S.C. Lee, D.W. Sim, J. Lee, K.Y. Jeong, K.H. Park, J.H. Lee, J.D. Kim, J.W. Park, Comparison between newly developed and commercial inhalant skin prick test reagents using *in vivo* and *in vitro* methods, *J. Korean Med. Sci.* 33 (13) (2018) e101.
- [4] J.L. Rice, G.B. Diette, C. Suarez-Cuervo, E.P. Brigham, S.Y. Lin, M. Ramanathan Jr., K.A. Robinson, A. Azar, Allergen-specific immunotherapy in the treatment of pediatric asthma: a systematic review, *Pediatrics* (2018), e20173833. <https://doi.org/10.1542/peds.2017-3833>.
- [5] G. D'Amato, L. Cecchi, S. Bonini, C. Nunes, I. Annesi-Maesano, H. Behrendt, G. Liccardi, T. Popov, P.J.A. Van Cauwenberge, Allergenic Pollen and Pollen Allergy in Europe, vol. 62(9), 2007 976–990.
- [6] T. Zuberbier, J. Lötval, S. Simoens, S. Subramanian, M.K.J.A. Church, Economic Burden of Inadequate Management of Allergic Diseases in the European Union: A GA2LEN Review, vol. 69(10), 2014 1275–1279.
- [7] A. Linneberg, K.D. Petersen, J. Hahn-Pedersen, E. Hammerby, N. Serup-Hansen, N.J.C. Boxall, M. Allergy, Burden of Allergic Respiratory Disease: A Systematic Review, vol. 14(1), 2016 12.
- [8] V. Cardona, P. Demoly, S. Dreborg, A.F. Kalpaklioglu, L. Klimek, A. Muraro, O. Pfaar, T.A. Popov, H.J.A. Hoffmann, Current Practice of Allergy Diagnosis and the Potential Impact of Regulation in Europe, vol. 73(2), 2018 323–327.
- [9] S. Vieths, S. Scheurer, J. Reindl, D. Luttkopf, A. Wangorsch, M. Kastner, T. Haase, D. Haustein, Optimized allergen extracts and recombinant allergens in diagnostic applications, *Allergy* 56 (Suppl. 67) (2001) 78–82.
- [10] D. Larenas-Linnemann, L.S. Cox, European allergen extract units and potency: review of available information, *Ann. Allergy Asthma Immunol.* 100 (2) (2008) 137–145.

- [11] M. Patterson, J.J.C. Slater, E. Allergy, Characterization and Comparison of Commercially Available German and American Cockroach Allergen Extracts 1, vol. 32(5), 2002 721–727.
- [12] R. Rossi, G. Monasterolo, G.J.E.A.A.C.I. Passalacqua, The Biological Potency of Different Extracts for Sublingual Immunotherapy Assessed by Skin Prick Tests, vol. 42 (3) (2010) 112–4.
- [13] K.Y. Jeong, C.S. Hong, J.S. Lee, J.W. Park, Optimization of allergen standardization, *Yonsei Med. J.* 52 (3) (2011) 393–400.
- [14] B. Jia, C.O. Jeon, High-throughput recombinant protein expression in *Escherichia coli*: current status and future perspectives, *Open Biol.* 6 (8) (2016).
- [15] G.L. Rosano, E.A. Ceccarelli, Recombinant protein expression in *Escherichia coli*: advances and challenges, *Front. Microbiol.* 5 (2014).
- [16] M. Merlin, E. Gecchele, S. Capaldi, M. Pezzotti, L. Avesani, Comparative evaluation of recombinant protein production in different biofactories: the green perspective, *Biomed. Res. Int.* 2014 (2014).
- [17] J.E. Cronan, *Escherichia coli* as an experimental organism, *eLS* (2014) <https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780470015902.a0002026.pub2>.
- [18] H.P. Sørensen, K.K. Mortensen, Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*, *Microb. Cell Factories* 4 (2005) 1.
- [19] T. San-Miguel, P. Pérez-Bermúdez, I. Gavidia, Production of soluble eukaryotic recombinant proteins in *E. coli* is favoured in early log-phase cultures induced at low temperature, *SpringerPlus* 2 (2013).
- [20] B. Muntari, A. Amid, M. Mel, M.S. Jami, H.M. Salleh, Recombinant bromelain production in *Escherichia coli*: process optimization in shake flask culture by response surface methodology, *AMB Express* 2 (2012) 12.
- [21] M. Fereidouni, M. Naghizadeh, N. Mahavar, M. Mousavi, A. Derakhshani, The most common aeroallergens among allergic rhinitis patients in Birjand city, Iran, *Allergy* (2018) 709 WILEY 111 RIVER ST, HOBOKEN 07030-5774, NJ USA.
- [22] M. Fereidouni, R. Farid Hossini, F. Jabbari Azad, M. Ali Assarezagadan, A. Varasteh, Skin prick test reactivity to common aeroallergens among allergic rhinitis patients in Iran, *Allergol. Immunopathol.* 37 (2) (2009) 73–79.
- [23] C. Vaquero, A. Rodriguez-Torres, J. Rojo, R. Perez-Badia, Airborne pollen of allergenic herb species in Toledo (Spain), *Environ. Monit. Assess.* 185 (1) (2013) 335–346.
- [24] Z. Wang, Y. Gao, Y. Zheng, Analysis on allergens in patients with allergic rhinitis in Urumqi Xinjiang, *Lin chuang er bi yan hou tou jing wai ke za zhi = J. Clin. Otorhinolaryngol. Head Neck surg.* 29 (15) (2015) 1328–1331.
- [25] L. Ferrer, J. Carnes, B. Rojas-Hijazo, M.A. Lopez-Matas, M.T. Sobrevia, C. Colas, Assessing degree of flowering implicates multiple *Chenopodiaceae/Amaranthaceae* species in allergy, *Int. Arch. Allergy Immunol.* 158 (1) (2012) 54–62.
- [26] H. Safarpour, S.B. Banadkoki, Z. Keshavarzi, M.H. Morowvat, M. Soleimanpour, S. Pourmolaei, F.H. Shirazi, Expression analysis and ATR-FTIR characterization of the secondary structure of recombinant human TNF- α from *Escherichia coli* SHuffle® T7 express and BL21 (DE3) cells, *Int. J. Biol. Macromol.* 99 (2017) 173–178.
- [27] H.R. Nouri, M. Sankian, F. Vahedi, D. Afsharzadeh, L. Rouzbeh, M. Moghadam, A. Varasteh, Diagnosis of *Chenopodium album* allergy with a cocktail of recombinant allergens as a tool for component-resolved diagnosis, *Mol. Biol. Rep.* 39 (3) (2012) 3169–3178.
- [28] M.M. Bradford, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal. Biochem.* 72 (1–2) (1976) 248–254.
- [29] H. Yang, S. Yang, J. Kong, A. Dong, S. Yu, Obtaining information about protein secondary structures in aqueous solution using Fourier transform IR spectroscopy, *Nat. Protoc.* 10 (3) (2015) 382–396.
- [30] J.E.P. Larsen, O. Lund, M. Nielsen, Improved method for predicting linear B-cell epitopes, *Immunome Res.* 2 (1) (2006) 2.
- [31] P.I. Haris, F. Severcan, FTIR spectroscopic characterization of protein structure in aqueous and non-aqueous media, *J. Mol. Catal. B Enzym.* 7 (1–4) (1999) 207–221.
- [32] Z. Movasaghi, S. Rehman, D.I. ur Rehman, Fourier transform infrared (FTIR) spectroscopy of biological tissues, *Appl. Spectrosc. Rev.* 43 (2) (2008) 134–179.
- [33] A. De Marco, Strategies for successful recombinant expression of disulfide bond-dependent proteins in *Escherichia coli*, *Microb. Cell Factories* 8 (1) (2009) 26.
- [34] M. Fathi-Roudsari, A. Akhavan-Tehrani, N. Maghsoudi, Comparison of three *Escherichia coli* strains in recombinant production of reteplase, *Avicenna J. Med. Biotechnol.* 8 (1) (2016) 16.
- [35] G. Ren, N. Ke, M. Berkmen, Use of the SHuffle strains in production of proteins, *Curr. Protoc. Protein Sci.* 85 (1) (2016) 5.26. 1–5.26. 21.
- [36] M. Nasiri, J. Babaie, S. Amiri, E. Azimi, S. Shamshiri, V. Khalaj, M. Golkar, P. Fard-Esfahani, SHuffle T7 strain is capable of producing high amount of recombinant human fibroblast growth factor-1 (rhFGF-1) with proper physicochemical and biological properties, *J. Biotechnol.* 259 (2017) 30–38.
- [37] J. Chen, J.L. Song, S. Zhang, Y. Wang, D.F. Cui, C.C. Wang, Chaperone activity of DsbC, *J. Biol. Chem.* 274 (28) (1999) 19601–19605.
- [38] C.P. Papaneophytou, G. Kontopidis, Statistical approaches to maximize recombinant protein expression in *Escherichia coli*: a general review, *Protein Expr. Purif.* 94 (2014) 22–32.
- [39] U. Reineke, R. Sabat, Antibody epitope mapping using SPOT™ peptide arrays, *Epitope Mapping Protocols*, Springer 2009, pp. 145–167.
- [40] K. Yoshimatsu, J. Arikawa, Antigenic properties of N protein of hantavirus, *Viruses* 6 (8) (2014) 3097–3109.
- [41] N. Salimi, W. Fleri, B. Peters, A. Sette, Design and utilization of epitope-based databases and predictive tools, *Immunogenetics* 62 (4) (2010) 185–196.