Application of a Seamless and Restriction Endonuclease-free Cloning Method to Produce Recombinant Full-length N-terminal His-tagged Streptolysin O in E. coli

Mohammad Hassan Kheirandish¹M.Sc., Hossein Zarei Jaliani¹, ²Ph.D.
Behnaz Rahmani¹M.Sc.

¹Department of Medical Genetics, Faculty of Medicine, Shahid Sadoughi University of Medical Sciences, Yazd, Iran.
²Department of Advanced Medical Sciences and Technologies, Faculty of Paramedicine, Shahid Sadoughi University of Medical Sciences, Yazd, Iran.

Article history
Received 28 Jun 2017
Accepted 12 Aug 2017
Available online 30 Aug 2017

Key words
Pore forming toxin
Seamless cloning
Streptolysin O

Background and Aims: DNA cloning, sub-cloning and site directed mutagenesis are the most common strategies in nearly all projects of recombinant protein production. The classical method of restriction site cloning is unsatisfactory due to the need for supply of restriction enzymes and the inefficiency of the digestion reaction. Many new methods, including recombinatorial cloning and ligation independent cloning need additional enzymes and kits. In this project we insert a full-length streptolysin O gene into an expression plasmid without using any uncommon commercial enzymes.

Materials and Methods: Streptolysin O gene was amplified by polymerase chain reaction (PCR) and introduced into the pPSG-IBA35 vector using a quick-change PCR. At the same time the gene was double digested and sub-cloned into pET28a (+). Both constructs were introduced into BL21 DE3 cell. Proteins were purified by Ni-NTA column and hemolytic activity was evaluated by spectrophotometry using human red blood cells.

Results: Streptolysin O was subcloned into the pET28a (+) and pPSG-IBA35 vectors and expressed in E. coli. Protein was purified with over 90% purity. The IC₅₀ of C and N terminus his-tagged protein were 0.22 and 0.29 µg/ml, respectively in hemolysis assay.

Conclusions: This study showed for the first time that full-length streptolysin O can be expressed in E. coli cytoplasm without any toxicity for the bacteria itself. The only additional amino acids expressed on the protein were his-tag. To study the role of this toxin it would be better to express the protein with the same strategy to have minimal extra amino acids on the protein.
Introduction

The most common class of protein toxins comprises family members with the ability to penetrate nearly all kind of biological membranes. They are called pore forming toxins. Elucidating their mode of action is at the beginning, but there is no doubt that all members are produced in a water-soluble form and oligomerize into multimeric assemblies inside the membranes after a more or less specific binding event [1, 2]. Two alpha and beta main subclasses are separated based on the structural properties. Other classes are cholesterol-dependent cytolysin (CDC) and small pore forming toxins [3, 4]. CDC monomers with molecular weight of 50-70 kDa are secreted mainly by gram positive bacteria and their assemblies in the membrane contain as many as 30-50 monomers, forming a large pore inside the lipid bilayer. CDCs require cholesterol as binding receptor in the membrane [5, 6]. Streptolysin O (SLO), a prototype member of the CDC family is produced in groups A, C and G of the hemolytic streptococci strains [7]. Like the other CDCs, it is activated by thiol-containing compounds and it is therefore an oxygen-sensitive toxin [8, 9]. Anti-sera against SLO has been used in diagnosis of the SLO producing bacteria [10, 11]. SLO has also shown potent anti-cancer properties [12, 13] and has been used to study macromolecular delivery [14-19]. There are a number of reports for the expression and purification of the SLO as a recombinant protein. GST-tag [20, 21], his-tag [20, 22], GST and his-tag [23] and maltose binding protein [24] have been used for the production of SLO as fusion protein. In all versions of the recombinant SLO reported in literature there are a large number of additional amino acids from the fusion tag or from the vector backbone [20-24]. In some studies, a number of amino acids have been removed from the N-term or C-term of the protein and the additional amino acids have been replaced [21]. In addition to the extra amino acids, it has been reported that SLO might have toxicity for the expression strains of E. coli and an N-terminal fragment should be removed from SLO for the efficient expression of the recombinant SLO production [20, 22]. Both of the low and high copy number cloning plasmids in which there must not be any effective expression of the target gene has failed to clone the SLO gene in E.coli [22]. Therefore, it has not yet been established whether SLO protein could be expressed as full-length and without any additional amino acids at the N-term or not. In this project, it was applied a seamless and restriction-free cloning method to introduce SLO gene into the pPSG-IBA35 vector so that the only amino acids added to the gene product will be a 6x his-tag.

Material and Methods

All chemicals were from molecular biology grade (Sigma or Merck Inc.).

Gene

Full-length coding sequence of streptolysin O (Uniprot Code: P0C0I3) gene, without the signal peptide was codon optimized and
chemically synthesized by Pishgaman Gene Transfer Co.

**C-terminal his-tag subcloning in pET28a**

The gene was received in a cloning vector (pUC57) and the construct pUC57-SLO was introduced into NovaBlue competent bacteria. One of the transformants was cultured and the construct was extracted and double digested with NcoI and XhoI restriction enzymes, and subcloned into pET28a (+) expression vector with a C-terminal hexa-histidine tag (His-tag) resulted in the expression cassette pET28a-SLO.

**N-terminal his-tag subcloning in pPSG-IBA35**

Purified pUC57-SLO construct was used as a template in a PCR reaction for the amplification of the SLO gene. Primer sequences were as follows;

F1: ATCACCATCACCAACCATGGTAAACAAACAGAACACCGCTTC
R1: CAAGCTTGCGGGTGGCTCCCCTCGAGTCA

Underlined nucleotides were complementary to the SLO gene sequence. Extra nucleotides at the 5’ of the primers were complementary to the backbone of the pPSG-IBA35 plasmid sequence. A stop codon was inserted in the middle of the reverse R1 primer (bold nucleotides) so that the inserted gene will have a stop codon after SLO sequence and the downstream sequences of the pPSG-IBA35 vector will not be added to the protein. In a second PCR reaction, a 300 bp PCR product was amplified using pPSG-IBA35 plasmid as template with the following primers;

F2: CAGCAACCGCGGCGCTTTTTAC
R2: GAAGCGGTGTCTCTGTGGTTACCATGGTGTTAGGTGATGGCTA

Underlined nucleotides in the R2 sequence were complementary to the first nucleotides of the SLO gene after the start codon. Two PCR products (300 bp and 1614 bp in length) were used as priming sequences in a SOEing PCR. The product of 1914 bp sequence length was used as megaprimer in a quick-change PCR to replace the insert sequence of the vector with SLO gene.

**SOEing PCR:** Two cleaned-up PCR products (300 bp and 1914 bp) were mixed and denatured for 1 min. 150 seconds of extension were applied for 6 cycles to extend the fragments to a 1914 bp longer fragment. Then, F2 and R1 primers were used to amplify the longer fragment (1914 bp) for 30 cycles (in the same reaction).

**Quick-change PCR:** Product of the SOEing PCR was used as megaprimer (1914 bp) to replace the MCS of the pPSG-IBA35 vector with the SLO gene sequence. 4 minutes of extension for 20 cycles were used in the EmeraldAmp® polymerase master mix to synthesize the plasmid backbone. Then, 20 minutes of pfu treatment was applied to remove the A-overhang of the products.

NovaBlue competent bacteria were transformed with the DpnI treated product of the above-mentioned quick-change PCR reaction. Transformants were checked for the presence of the SLO gene using a colony-PCR with the following primers; F1 and R1. PCR in all steps was performed using a Labnet MultiGene™ Optimax Thermal Cycler.
Expression induction and purification

Competent *E. coli* BL21 (DE3) was used for the transformation and production of recombinant SLO protein. Briefly, a colony of bacteria harboring the pET28a-SLO or the pPSG-SLO construct was cultured in lysogeny broth (LB) medium containing 50 µg/ml of kanamycin (or ampicillin for the pPSG-SLO construct) at 37°C. Fresh aliquots of this culture media were used for inoculation of another 100 ml LB medium. The culture was shaken (150 rpm) to obtain the OD$_{600}$ nm of approximately 0.7 and induced by the addition of 0.5 mM isopropyl β-D-1-thiogalactopyranoside. After 4 hours of induction, the bacteria were harvested by centrifugation at 6,000 rpm for 10 min. Bacterial pellet was bead beaten 15 seconds for 10 times. At each interval pellet was incubated on ice for at least 30 sec. Lysis buffer contained 50 mM Tris HCl pH=8.0, 300 mM NaCl, 10% Glycerol, and 3 mM dithiothreitol. Soluble fraction was centrifuged (12,000 rpm at 4°C) and the supernatant was loaded onto a Ni-NTA column (Qiagen Inc.) to purify the His-tagged SLO. Elution Buffer is same as lysis buffer, except for the 300 mM NaCl and 300 mM imidazole. The eluted fractions were analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) for the presence of protein, and fractions contained the SLO protein were dialyzed against the imidazole free buffer.

Hemolysis Activity

To examine the hemolytic activity of purified SLO [25], volunteer human red blood cells (RBCs) were washed thrice with phosphate buffered saline PBS, pH 8.0. RBCs absorbance at 600 nm was adjusted to around 1.0. Changes in turbidity (A$_{600}$ nm) were recorded in a Unico UV2100 spectrophotometer after a 15 min. incubation of diluted RBC solution with different concentrations of the purified SLO toxin. Hemolytic activity is expressed in percent using the following formula: 
\[
\text{[H(\%)]=}(A_{\text{max}}-A_{\text{obs}})/ (A_{\text{max}}-A_{\text{min}}) \times 100
\]
where the $A_{\text{max}}$ and the $A_{\text{min}}$ represents the absorbance of the intact and completely lysed RBC respectively. $A_{\text{obs}}$ is the absorbance of each sample recorded after 15 min. incubation with each dilution of the toxin. Spectrophotometry in all steps of the present study was performed using a Unico™ 2100 spectrophotometer.

Results

Sub-cloning in pET28a and pPSG-IBA35

Codon-optimized open reading frames (ORF) of the SLO protein (Uniprot Code: P0C0I3) except of the signal sequence (total 538 aa) was ordered in pUC57-amp vector. After cloning in NovaBlue competent bacteria and plasmid extraction and purification, the pUC57-SLO construct and pET28a (+) were double digested with NcoI and XhoI restriction enzymes. Gel purified SLO fragment was ligated into pET28a and the same competent bacteria was used for transformation of pET28a-SLO constructs (data not shown).

This subcloning of the SLO gene has led to the C-terminal his-tag labeling of SLO protein. Using NcoI and XhoI restriction sites of the pET28a expression system will add two amino acids (Leu-Glu) to the C-terminal of the protein after which the his-tag will expose. So
only eight amino acids have been added to the SLO protein. In the second step, for the subcloning of the SLO gene with N-terminal his-tag, pPSG-IBA35 vector was used for the reason that the N-terminal his-tag in this expression system (StarGate™) is fused to the gene without additional undesired sequences present in pET expression system (Fig. 1; steps A, B, C and D). As illustrated in Fig. 1 (step B), two independent PCR experiments were done using primers mentioned in before. As seen in Fig. 2A, sizes of the PCR products were 300 and 1614 bp as predicted. Two fragments were fused using a SOEing PCR and the 1914 bp PCR product was produced for using as megaprimer in quick-change PCR (Fig. 2-B). The study was approved by the Ethics Committee of Shahid Sadoughi university of medical sciences, Yazd, Iran.

![Fig. 1. Seamless cloning strategy. For subcloning of the SLO gene into expression vector with N-terminal his-tag, one can add sequences complementary to the backbone of the expression vector to the target gene using a simple PCR. Primers in this PCR have extra sequences at their 5' ends (R1 and R2) complementary to the backbone of the expression vector (A). After this step, it is possible to directly synthesize the vector in a quick-change PCR using the product of the A step as a megaprimer. Two complementary regions should be long enough to have a Tm around 50-60°C (D). In the other format which was done in this study, it is possible to add some parts of the expression vector to the target gene before doing a quick-change PCR. Two independent PCR reactions with overlapping sequences between the two primers (R2) are done (B). SOEing PCR using this overlapping sequence (R2) and two flanking sequences complementary to the backbone of the expression vector (R1 and R3 in C step) will lead to PCR product which is the megaprimer for the quick-change PCR (B).](image-url)
Megaprimer (1914 bp fragment) which is synthesized by SOEing PCR was cleaned up and used in a Quick-change PCR using EmeraldAmp® polymerase. The product of the quick-change was seen as a big clear band on 1% agarose gel (Fig.3). Megaprimer in this quick-change method is composed of SLO gene and a fragment of pPSG-IBA35 vector. A product of the Quick-change PCR was DpnI treated to remove the template plasmids and transformed into NovaBlue chemically competent cells. One of the transformants was picked up, and its plasmid was sequenced with universal primers for the vector backbone. No mutation was seen in the whole gene (data not shown).
Recombinant his-tagged SLO expression induction and purification

The pET28a-SLO plasmid was extracted and purified from NovaBlue and was transformed into chemically competent BL21 (DE3) cells. Induced cultures of transformants were used for the purification of the C-terminal his-tag SLO recombinant protein by Ni-NTA affinity resin as seen in Fig. 4. Recombinant SLO protein was induced in BL21 DE3 cells in both C and N-terminal his-tag forms as about 61 kDa protein band on a 10% SDS-PAGE. 100 ml cultures of the BL21 bacteria containing two forms of recombinant his-tagged SLO construct were induced overnight; after harvesting cell pellet and sonication on ice, soup of the bacteria was applied on a Ni-NTA resin. As seen in Fig. 5 and Fig. 6, recombinant SLO protein can be expressed and purified as C-terminal and also N-terminal his-tagged fusion protein, respectively.

![Fig. 4. Expression induction of the SLO gene. SLO protein was induced by 0.5 mM IPTG in pET28a as C-terminal his-tag fusion and pPSG-IBA35 as N-terminal his-tag fusion protein. Lane 1 and 2; BL21 cells containing pET28a-SLO and pPSG-IBA35-SLO constructs respectively before induction. Lane 3 and 4; BL21 cells containing pPSG-IBA-SLO construct induced by adding 0.5 mM IPTG to two selected clones. Lane 5 and 6; BL21 cells containing pET28a-SLO construct induced by adding 0.5 mM IPTG to two selected clones. Lane 7; prestained protein size marker (Sinaclon cat no. PR911641).](image)

![Fig. 5. C-terminal his-tagged SLO purification. Lane 1; soup of the induced bacteria before column, lane 2; flowthrough, lane 3; protein size marker (CMG Co.), lane 4; wash, lane 5; first fraction of the elution.](image)
Hemolytic Activity

Different concentrations of the purified toxin were incubated with diluted human RBC to evaluate the hemolytic activity of SLO. The IC50 of N-term and C-term his-tagged forms of SLO toxin for the human red blood cells were calculated to be 0.29 and 0.22 µg/ml, respectively (Fig. 7).

Fig. 7. SLO hemolytic activity. N-terminal his-tagged SLO was applied to 5% human red blood cells suspension in a hemolysis assay. SLO serial 1/2 dilutions were incubated with human RBC at 37°C for 15 min and RBC lysis was monitored by recording the absorbance at 600 nm. Hemolytic activity is expressed in percent using the following formula: \[ H(\%) = \left( \frac{A_{\text{max}} - A_{\text{obs}}}{A_{\text{max}} - A_{\text{min}}} \right) \times 100 \] where the \( A_{\text{max}} \) and the \( A_{\text{min}} \) represents the absorbance of the intact and completely lysed RBC, respectively. \( A_{\text{obs}} \) is the absorbance of each sample recorded after 15 min incubation with each dilution of the toxin.

Fig. 6. N-terminal his-tagged SLO purification. Lane 1: soup of the induced bacteria before column, lane 2: flowthrough, lane 3: first drops of the wash, lane 4: protein size marker (CMG Co.), lane 5: last drops of the wash, lane 6: first fraction of the elution.
Discussion

SLO is secreted in nearly all strains of group A *streptococcus* bacteria, a common gram-positive pathogen responsible for different human diseases. SLO exerts its effects at several stages and its role in the process of infectivity has to be studied in more details. SLO has been expressed and purified as a recombinant protein by a few researchers. There is still controversy about the possibility of SLO production as full-length protein in *E. coli*. In some studies, it has been expressed as fusion protein with GST-tag [21] or maltose binding protein [24]. Weller et al. produced a truncated form of SLO (496 aa) as fusion with maltose-binding protein. They reported that the uncleaved fusion protein (MBP-SLO) has hemolytic and cytotoxic activities comparable to the native SLO protein [24]. In the case of GST-tag fusion protein, when the mature full-length SLO protein has been tried, Kimoto et al. have not observed any recombinant protein production [20]. Velázquez et al. also attempted to express and purify the SLO as fusion with GST-tag. Their results showed that when the fusion was not full-length (amino acid 78-571 of the SLO protein), the recombinant protein could be expressed and purified well, but in a low yield of 1.5 mg/liter [21]. Even it has been reported that the cloning vectors carrying the mature full-length SLO gene without any fusion gene could not be amplified in the *E. coli* cytoplasm [22].

Previous studies have failed to address the production of mature SLO protein with minimal extra amino acids of the vector backbone or the fusion protein itself. In some studies, it has been argued as unpublished results that the full-length SLO protein could not be expressed in *E. coli* even as the fusion with GST-tag [20]. Regarding these controversies we attempted in this report to produce SLO protein as mature and full-length protein without any large fusion tag. The only extra amino acids added to the SLO protein in the present study was the hexahistidine tag in either C or N-terminal of the SLO protein.

For the C-terminal his-tag fusion protein expression, pET28a (+) expression vector was used. NcoI and XhoI restriction sites of the multiple cloning site of this vector removed all unnecessary codons except of the XhoI site itself and a C-terminal his-tag. Although, for the N-terminal his-tag expression of the recombinant protein, the majority of the commercial vectors (e.g. pET series) have many additional codons, which was not necessary in our study. So, it was decided to use an alternative strategy to produce SLO protein with only an N-terminal hexa his-tag. Two pairs of forward and reverse primers were designed. The first one amplified the SLO gene (a 1614 bp product) with additional nucleotides, which was complementary to the pPSG-IBA35 vector. The second set of primers amplified a region of the pPSG-IBA35 vector (a 300 bp product) upstream to the start site of the target gene. These two products were used in a SOEing PCR reaction to produce a 1914 bp fragment containing SLO gene fused to a fragment of the vector. This really long megaprimer was used in a quick-change PCR reaction with a
long-PCR polymerase enzyme to produce recombinant construct. In the present study, using long-PCR polymerase enzyme in quick-change PCR has led to higher yields of quick-change step. Most researchers use pfu, Phusion, and other proofreading enzymes to avoid deleterious mutations in the gene or in the most important regions of the vector, especially when amplifying the whole plasmid in an inverse PCR protocol [26-29].

In this method, megaprimer of the quick-change step was the whole insert gene plus 300 bp upstream of the start codon. So, the only regions polymerized in the quick-change step were the remaining regions of the plasmid backbone (selectable marker, origin and etc.).

SOEing PCR (Gene Splicing by Overlap Extension) can be used to join two and even more fragments of DNA and produce a larger DNA fragment. The difficulty of the long fragments manipulation (1-2 kb) has been addressed before [30-33]. We used a SOEing PCR method to produce a 1914 bp fragment containing SLO gene and a fragment of the expression vector upstream of the gene start site. The advantage of this was that the upstream region of the strat site is high-risk (promoter, RBS and etc.) and amplification of this region before starting the quick-change PCR reaction can minimize the risk of deleterious mutations. In fact, the only regions to be polymerized in the quick-change step are the low-risk regions of the plasmid backbone (selectable marker, origin and etc.).

Megaprimering has been extensively discussed in detail elsewhere. Briefly, a primary PCR reaction should be done and its product is used in the second PCR reaction as a primer [34, 35]. If the second PCR is a kind of inverse-PCR (like the quick-change PCR in mutagenesis protocols), the PCR product of the first reaction is sufficient as both 3'-OH ends of the double stranded PCR product can prime the second reaction. Megaprimering by means of only one mutagenic primer and a universal primer on the vector backbone has also been addressed [36].

**Conclusion**

In this study, it is for the first time that SLO is expressed without any additional amino acids (except that of 6 histidine residues at the N or C-terminal) and that the whole sequence of the mature protein (538 aa by removing the signal sequence) is expressed in *E. coli* without any toxicity for the bacteria. To study the role of this toxin in the infection process, it would be better to express the recombinant SLO protein with the same strategy to have minimal extra amino acids or tags on the protein.

**Conflict of Interest**

There is no conflict of interest to declare.

**Acknowledgements**

This work was supported by a grant from the deputy of research and technology, Shahid Sadoughi university of medical sciences and was written based on a data set of a medical biotechnology M.Sc. thesis registered in faculty of medicine, Shahid Sadoughi university of medical sciences.
References


