Synthesis, Subcloning, and Sequencing of a Codon Optimized Variant of Proteinase Inhibitor II Designed For Expression in *E. coli*

Arshia Alimohammadi, Jacqueline Siu, Alan Stachowiak, Emily Fogarty

Department of Microbiology and Immunology, University of British Columbia

Abstract

Proteinase Inhibitor II (PI2) is a potato tuber peptidase with many potential applications including pest control, radiation protection, and hunger suppression. Recombinant production using *E. coli* is a cost-effective approach to produce PI2. In this study, we re-designed the previously used *pi2* gene for *E. coli* expression and subcloned it into pCR2.1-TOPO-TA vector after PCR amplification. The re-designed *pi2* gene fragment is based on the native cDNA sequence that was codon optimized for *E. coli* expression. The nucleotide sequence of the synthetic *pi2* gene was verified by sequencing. The subcloned *pi2* in plasmid pCR2.1-TOPO-TA-PI2 is flanked by several restriction enzyme cut sites and can be easily cloned into protein expression vectors such as pET32 in future studies.

INTRODUCTION

Proteins perform a vast array of functions within living organisms and have been harnessed to serve as pharmaceuticals and molecular tools (1). Proteins produced by eukaryotic cells can be difficult to express in large concentrations (1). In addition, the use of eukaryotic cells, particularly multicellular organisms, for protein production may be complex, expensive and/or unethical. To circumvent these issues, eukaryotic DNA is often transformed into prokaryotic cells (most commonly *Escherichia coli*) and expressed as recombinant protein. However, to express eukaryotic genetic material in prokaryotic cells, GC content and codon frequency need to be
optimized since these genetic parameters can be significantly different between prokaryotes and eukaryotes (2). The
frequency of the codon usage usually reflects the abundance of their cognate tRNAs, therefore, an infrequently used
codon is not translated at a high level, resulting in low observed protein production (2).

A common obstacle to efficient protein production in prokaryotic systems is the improper folding of proteins resulting in insoluble aggregates known as inclusion bodies (3). The purification and refolding of the aggregated proteins is characterized by poor recovery, many repeated experiments needed to optimize the recovery and accounts for the majority of the costs associated with the production of recombinant proteins in E. coli (4).

Often, the ideal solution involves modifying the protein or expression environment for the protein to be soluble.

Thioredoxin is a small protein that facilitates the reduction of other proteins through cysteine thiol-disulfide exchange (5). The linkage of E. coli thioredoxin (TrxA) to heterologous protein has been shown to decrease inclusion body formation (6); however, it is unclear how thioredoxin helps with solubility (6) or if it acts as a chaperone (7) when it aids in the production, and folding of proteins.

Proteinase Inhibitor II (PI2) is a 21 kDa dimeric, cysteine-rich, heat stable endo-acting peptidase that inhibits chymotrypsin and trypsin (8, 9). It is found in potato tubers and could be used as a pest controlling agent by inhibiting amino acid absorption of herbivores and pathogens by inhibiting their digestive enzymes (10). PI2 could also be used for radiation and UV damage protection by inhibiting proteases involved in the propagation of cellular damage (11, 12). Medically, it could be used as a hunger suppressant as it has been shown to raise cholecystokinin levels, leading to satiety and decreased caloric intake (13).

PI2 is composed of 154 amino acids, which include 16 cysteine residues (9). The 16 cysteine residues in its structure form 8 disulfide bonds (14, 9). Since it has many disulfide bonds, PI2 is predicted to be difficult to express and to form insoluble inclusion bodies in E. coli, as prokaryotic cells do not have the endoplasmic reticulum’s oxidizing environment to promote the formation of these bonds (9, 14).

Previous work has attempted to generate PI2 linked to TrxA in pET32a(+) plasmid constructs, however, either a PI2 sequence containing an intron was inserted into the plasmid (15) or PI2 was absent in these putative clones (16). This study revisited the initial phases of the PI2 project. We described the analysis and design of an intron-free, codon optimized pi2 sequence with appropriate GC content for expression in E. coli.

MATERIALS AND METHODS
Analyzing previous PI2 sequence

NCBI Standard Nucleotide BLAST was used to compare Przeworski et al.’s JEMI PI2 sequence with Keil et al.’s primary nucleotide and cDNA sequence (16, 17). GenScript’s Rare Codon Analysis Tool was used to determine the GC content and how frequently each PI2 codon is found in native E. coli codon usage patterns. This tool can be found on the Genescript webpage (http://www.genscript.com/cgi-bin/tools/rare_codon_analysis).

Designing PI2 gBlock gene fragment

Keil et al.’s PI2 cDNA sequence was codon optimized for E. coli expression using OPTIMIZER (18). GC content and E. coli codon usage frequency was verified as stated above. NcoI and EagI restriction enzyme cut sites were added to either end of the gene after 6 - 7 non-specific nucleotides in from the free ends of the DNA insert; these restriction sites are separated by more than 12 bp in the pET32a/pET30b multiple cloning sites (19, 20).

Strains and plasmids

pET30b(+), pET32a(+), supplied by Novagen, were used for protein expression and TOPO-TA, supplied by Invitrogen, was used as a storage vector. E. coli DH5α strain, supplied by the microbiology department of University of British Columbia, was used for all transformations and plasmid isolations. Cells containing each of the plasmids were grown overnight at 37°C at 180 RPM in Lysogeny Broth (LB) liquid medium (pH 7.0) containing tryptone (1.0%), yeast extract (0.5%) and NaCl (1.0%).

Plasmid isolation

Plasmid DNA was isolated from DH5α E. coli using the Invitrogen PureLink HQ Mini Plasmid DNA Purification Kit (Cat# K2100-01). DNA concentration and purity was measured via 260nm and 280nm absorbances using a ThermoScientific NanoDrop 2000c Spectrophotometer.

Preparing competent DH5α E. coli

An isolated colony was grown overnight in 1 ml of LB broth and added to 99 ml of LB broth for a culture volume of 100 ml. Culture was shaken at 180 RPM at 37°C until an OD of 0.8-0.9 was achieved. All centrifugation steps are performed at 4000 x g for 10 minutes. Suspended cultures were pre-cooled and centrifuged. Pellets were
resuspended in 100 ml of cold dH₂O and centrifuged. Pellets were resuspended in 50 ml of cold dH₂O and centrifuged. Pellets were resuspended in 20 ml of cold sterile 10% glycerol and centrifuged. Finally, pellets were resuspended in 0.5 ml of cold sterile 10% glycerol. 40 µl of resuspension was aliquoted into each pre-cooled tube, and flash-frozen in liquid nitrogen. Competent cells were stored at -80°C.

Electroporation

Competent DH5α E. coli cells (40 µl) were thawed and transferred into plastic cuvettes with < 10 ng of isolated DNA. Using a BioRad MicroPulser (2.5 V, 200 ohm, 500 uF), cells were transformed with DNA and added to 1 ml LB and incubated at 37°C at 350 RPM for 60 minutes. After incubation, cells were spread-plated on antibiotic selective LB plates and incubated overnight at 37°C. Colonies were counted the following day.

Restriction digest of PI2 gBlock and vectors

100 ng of PI2 was digested with 1µl of NcoI and 1µl EagI (NEB Cat# R3193, R350) with 3µl of 10X NEBuffer 3.1 (Cat# B7203S) and brought up to 30µl with dH₂O. Vectors pET30b(+) and pET32a(+) were subject to both single and double digests using the same reaction conditions, and bringing all reaction volumes 30µl with dH₂O. All reactions were incubated at 37°C for 60 minutes.

Ligation of PI2

Restriction digest samples were incubated at 60°C for 20 minutes to heat-inactivate enzymes. For ligation, 50ng of linearized vector was mixed with 5X molar excess of gBlock gene fragment with 1 μL of NEB Quick DNA Ligase and 10 μL 2X Quick Ligase Buffer (Cat# M2200S). Samples were incubated at room temperature for 5 minutes and later used to transform competent DH5α.

PCR amplification of PI2

The following oligonucleotide primers were used to amplify the pi2 gene: forward primer (5’-CATCCATGGCTATGGACGTT-3’), reverse primer (5’-CTTGCCGCGGCATTATTAC-3’). All PCR reactions were carried out in volumes of 25 µl. Each reaction contained 2.5 µl of 10X polymerase buffer, 2.5 µl of 50 mM MgCl₂, 0.5 µl of 10 mM dNTPs, 0.5 µl of 10 uM forward and reverse primers, 0.2 µl of Platinum Taq DNA
polymerase (Invitrogen, Cat# 10966018), 1 µl of 0.1 - 1 ng template DNA and PCR grade dH₂O. PCR cycle conditions used are: an initial denaturation at 94°C for 2 minutes, 35 cycles of denaturation 94°C for 30 seconds, annealing at 55°C for 30 seconds and extension at 72°C for 30 seconds, with a final extension at 72°C for 5 minutes. For colony PCR, colonies of interest were gently scraped off plates with a pipet tip, added to the appropriate reaction tube and denatured initially for an additional 8 minutes at 94°C to ensure cell lysis and to allow the release of template DNA. All PCR products were subsequently resolved on 1.5% agarose gel at 120V for 60 minutes and visualized by exposure to UV light. A Log-2 (0.1-10kb) DNA marker from NEB (Cat# 3200S) was used to determine band sizes.

**Isolation of PI2 PCR product**

Amplified PI2, which presented as a distinct band around 500bp, was cut out of the 1.5% agarose gel over a UV-box and purified using an Invitrogen PureLink Quick Gel Extraction Kit (Cat# K2100-12) according to the manufacturer’s instructions. The product of the gel extraction was used as the template for subsequent PCR reactions and insertion into various vectors.

**Insertion of PI2 into TOPO-TA vector**

In order to store PI2 for future use, a TOPO-TA Cloning Kit for Sequencing, with One Shot TOP10 Chemically Competent kit (Invitrogen, Cat# K4575-01) was used. Following the manufacturer's instructions, PI2 was inserted into TOPO-TA and TOP10 competent cells were transformed using heat shock. Transformed cells were plated on Kan LB plates supplemented with X-gal for blue/white screening.

**DNA sequencing of TOPO-TA+PI2 vector**

Sanger sequencing was done to sequence colonies with potential inserts. Isolated colonies were grown overnight to 5 ml and subject to DNA extraction using Invitrogen PureLink HQ Mini Plasmid DNA Purification Kit (Cat# K2100-01). Sample concentration was confirmed to be around 200 ng/µl using a ThermoScientific NanoDrop 2000c Spectrophotometer, as requested by NAPS (the sequencing facility). It was requested that the M13 forward primer be used for Sanger sequencing, and samples were delivered to NAPS on the UBC campus.
RESULTS

Analysis and optimization of pi2 nucleotide sequence for expression in E. coli

We first analyzed the pi2 sequences reported by Przeworski et al. and Keil et al. (16, 17). As shown in Figure 1, Przeworski et al. reported an open reading frame (orf) of 579 nucleotides while Keil et al. reported a 462 nucleotides orf. Alignment of these sequences using BLAST (basic local alignment search tool) to compare the potential pi2 sequence, we show that the JEMI pi2 gene sequence reported by Przeworski et al. (hereon referred to as JEMI-PI2) match the primary DNA sequence reported by Keil et al. (hereon referred to as Keil DNA), but not the cDNA sequence (hereon referred to as Keil cDNA) (Fig 1a, S1) (17). Keil cDNA sequence consists of two exons in different reading frames: 1- 53 bp and 171 bp - 468 bp. We conclude JEMI-PI2 sequence includes a 117 base pair intron.

In addition, based on a study reported by Sharp et al. describing E. coli codon usage, we determined that pi2 contains 53 codons used less than 50% of the time in E. coli (Fig 1c) (2). We also found that the GC content of pi2 sequence ranged from 10% - 63% (Fig 1d). GC content affects DNA stability and mRNA secondary structure; thus, GC content may contribute to codon bias. Given the intron, the poor E. coli codon bias, and non-optimal GC content, we decided to synthesize a new pi2 sequence rather than to do site-directed mutagenesis on dozens of sites. Therefore, we re-designed the pi2 gene to accommodate its expression in E. coli. The resulting construct does not contain an intron, all 155 codons are commonly found in E. coli, and the GC content is within the accepted E. coli range (Fig 1c, d) (2).

PCR amplification and cloning of pi2 into pCR2.1-TOPO-TA

Our attempts at directly cloning the pi2 fragment into a pET32 expression vector were unsuccessful. To ensure that the gene block wasn’t depleted, we decided to amplify the pi2 gBlock DNA using PCR. Primers were designed to PCR amplify the 500 bp sequence. As shown in Figure 2a, the gene block was amplified as evidenced by distinct bands at 500 bp. Fainter bands are visible at 1000 bp at all different template DNA concentration used. A water template control was included as a negative control (Fig 2a).

Following this, we cloned the pi2 PCR product into pCR2.1-TOPO using topoisomerase I. Single deoxyadenosine residues added to ends of pi2 PCR product allow insertion into pCR2.1-TOPO (21). After pCR2.1-TOPO pi2 ligation and electroporation into DH5α E. coli, blue/white colony screening on X-Gal and selection
on a kanamycin agar plate was used to identify colonies carrying plasmids with the pi2 insertion. Five white colonies (suspected pi2 positive colonies), and one blue colony were screened by PCR using the pi2 primers described above. PCR products were resolved by gel electrophoresis (Fig 2b). All five white colonies have a band at 500 bp. The 500 bp band is not seen in PCR performed on the blue colony. The water control also has a faint band at 500 bp, possibly a result of pi2 template contaminate from another sample. Two of the white colonies have 100 bp bands, possibly a result of primer dimerization. pi2 primers have 5 bases that could base pair resulting in primer dimers (22). However, it is more likely the band at 100 bp is contamination as it is only seen in colonies 1 and 2, while primer dimers should be seen in every lane. These results suggest that pi2 has been inserted into pCR2.1-TOPO plasmid.

Sequencing results show pi2 gene inserted into pCR2.1-TOPO

DNA sequencing was performed in order to confirm that the redesigned pi2 gene fragment had been cloned into pCR2.1-TOPO DNA. Analysis of the results sequence data indicated that the pi2 fragment had been inserted into pCR2.1-TOPO with the restriction enzyme site NcoI located at the 5’ of the inserted pi2 gene and EcoRI located at both ends of the inserted gene. Of the five colonies screened and sequenced, only plasmid pCR2.1-TOPO-TA-1-lenta_M13R had no mutations. The other four sequenced plasmids each had the same silent point mutation at the 81st nucleotide where the codon changed from CCA to CCG. The sequencing results for plasmid pCR2.1-TOPO-1-lenta_M13R are shown in Figure 5S.

DISCUSSION

For PI2 expression in E. coli, the proper primary amino acid sequence must be translated. For the majority of eukaryotes, the mRNA transcribed in the nucleus must be modified prior to its translation in the cytoplasm; modification may include intron removal and mRNA splicing (23). On the other hand, most prokaryotes including E. coli do not have introns or intron removal mechanisms. We determined that the JEMI-PI2 nucleotide sequence matched the eukaryotic potato pi2 gene, which contains a 117 base pair intron (Fig 1a, S1) (16, 17). The presence of an intron in JEMI-PI2 is problematic because E. coli is unable to splice the intron out, and the wrong mRNA sequence is used for translation. Not only are additional amino acids translated, but the second exon is out of frame with the first intron.
In addition to the intron in JEMI-PI2, the nucleotide sequence is not optimal for expression in *E. coli* due to a number of codons that are rarely used in *E. coli*. Although parts of the genetic code are degenerate, some species have different preferences for certain amino acids codons (2). As shown in Figure 1c, 30% of JEMI-PI2’s codons are preferentially used less than 50% of the time in *E. coli*. The literature suggests that it is important to account for codon bias when expressing recombinant proteins because if the distribution of codons is significantly different than the typical *E. coli* codon distribution then a reduction in quality or quantity of the heterologous protein could occur (24). However, there is evidence to suggest that even though different codons code for the same amino acid, there is still a difference in protein solubility and functionality (25). Thus, while it is important to optimize the codons for *E. coli* codon usage patterns, it is crucial to take into account these changes for downstream processes. In DNA, the guanine and cytosine nucleotide base pairs form a stronger bond with three hydrogen bonds while the adenine and thymine base pairs have a weaker bond with only two hydrogen bonds. Overall, the GC content affects the stability of the DNA and the secondary structure of the mRNA. For *E. coli*, it appears most native genes GC content to be between 15% and 70% (26). GC content is believed to help drive codon bias (26). As shown in Figure 1d, the GC content of *pi2* ranged from 10% to 63%. Given the presence of an intron, the poor *E. coli* codon bias, and highly varied GC content, we decided to design and order a new *pi2* sequence rather than do site-directed mutagenesis on dozens of sites. The designed *pi2* sequence is based on the published cDNA sequence, and we further optimized the codons for *E. coli* translation and ensured the GC content is within an acceptable range as shown in Figure 1b, c, d.

As shown in Figure 2a, *pi2* gBlock DNA was amplified using PCR. As expected there is the gene of interest at 500 bp, unexpectedly there are faint bands visible at 1000 bp for each lane. These 1000 bp bands are most likely a result of PCR product dimerization. Interestingly, when this experiment was repeated with samples placed in the middle of the thermocycler rather than the edges, we did not observe the 1000 bp bands (data not shown). For future PCR experiments, we suggest only placing samples in the middle wells due to unequal heating. We attempted to clone the *pi2* gBlock into pET30b and pET32a (Fig S2). Although we did not obtain the desired clone, we were able to verify that our DH5α *E. coli* were electrocompetent, and that the individual EagI and NcoI restriction enzymes were active (Fig S3).

We were unable to demonstrate successful ligation (Table 2). Issues arose with ligation when no growth was observed from any transformed cells. Likely, the problem was a non-functional ligase as neither samples nor single digest controls were seen to circularize as expected. Growth that should have been observed on ampicillin and
kanamycin plates (pET32a and pET30b respectively) was not present as can be seen from Table 2S. A non-functional ligase could be the result of ATP degradation in buffer during freeze-thaw cycles. Figure 3S shows that restriction enzymes EagI and NcoI were functional when used for single cuts. EagI is known to have low activity in a double digest in Tango buffer, therefore it is possible that when using EagI in conjunction with NcoI to digest pET30b/pET32a, EagI is rendered less efficient or not functional (27). The smallest band on the ladder is 100 bp, thus the 50 bp band excised out during digest, assumed present, could not be resolved. In the event that ligation and transformations were successful, it seems plausible that the PI2 product produced in *E. coli* could be toxic to the cells through the formation of inclusions bodies (28) or inhibition of proteases. However, this was not possible in our experiment as DH5α cannot express genes from pET30b or pET32a since it lacks the T7 polymerase required to drive expression from the T7 promoter sequences on these vectors.

In order to preserve the pi2 gBlock DNA for future use, while avoiding using ligase or restriction enzymes, pi2 was inserted into pCR2.1-TOPO. We used pCR2.1-TOPO as a storage vector, to replicate pi2 in *E. coli*. Figure 4 illustrates the cloning strategy for pi2. In the future pi2 DNA can be excised using restriction enzymes or amplified using PCR from pCR2.1-TOPO and ligated into another plasmid for protein expression.

**FUTURE DIRECTIONS**

The insertion of PI2 into pCR2.1-TOPO allows a future group to subclone PI2 into an expression vector (Fig 3). Although the restriction enzyme cut sites designed in the PI2 gBlock were EagI and NcoI, using another enzyme in the place of EagI is recommended, as EagI has lower activity in a double digest. An EcoRI site could be added into the PI2 using a primer designed with this site. Once the PI2 is isolated from pCR2.1-TOPO, it can be ligated directly into pET30b and pET32a.

After ligation into pET30b and pET32a, electrotransformation into *E. coli* strain DH5α could be used as a subcloning host. Once recovered, an *E. coli* strain that can express proteins from pET30b and pET32a, for example Origami, could be transformed with pET30b(+)PI2 and pET32a(+)PI2. Origami has a favourable environment for folding proteins with many disulfide bonds. The expression levels of PI2 could be determined by running whole cell lysates on SDS-PAGE. If expression cannot be determined by whole cell lysates, a His-tag purification would reduce background caused by cell debris and allow visualization of PI2 on a gel.
ACKNOWLEDGEMENTS

This study was facilitated by the University of British Columbia Microbiology and Immunology department. We would like to thank Dr. David Oliver and Chris Deeg for their knowledge and support throughout the course of this study. We would also like to thank the students previously involved with this project for providing us with the foundation of our study.

FIGURE LEGENDS

FIG. 1. Analyzing previous PI2 sequence and designing new PI2 gBlock.

(A) Representation of BLAST results comparing Keil et al.’s primary DNA and cDNA sequence with Przeworski et al.’s JEMI PI2 sequence, showing the presence of an intron, represented as a line, between 54 bp and 171 bp (B) Comparing Keil et al.’s cDNA sequence with AJAE-PI2’s E. coli codon optimized PI2 DNA sequence. The associated single letter amino acid codes are indicated above, vertical lines mark the open reading frame, and (+) indicates which nucleotide base are changed. (C) Histogram of how frequently PI2 codons are used (%) relative to E. coli codon usage patterns. (D) GC content % over JEMI-PI2 or AJAE-PI2 DNA sequence.

FIG. 2. PI2 PCR products visualized via gel electrophoresis in 1.5% agarose. (A) PCR amplification of PI2 gDNA. (B) Colony PCR of pCR2.1-TOPO(+)PI2 containing colonies to screen for PI2 inserts. 2-log ladder from NEB.

FIG. 3. Cloning representation of PI2 into pCR2.1-TOPO and attempted pET30b/pET32a insertion. (A) Plasmid map of pCR2.1-TOPO(+)PI2 designed in APE program illustrating enzyme cut sites and PI2 gBlock insertion. (B) Pictorial representation of PI2 cloning. Dashed arrow shows attempts to clone PI2 directly into
pET30b and pET32a. Solid black arrow represent successful insertion of PI2 into pCR2.1-TOPO. Grey arrows indicates future digests and ligations to insert PI2 into pET30b and pET32a. Plasmid map designed in APE program.

SUPPLEMENTARY FIG. 1. Comparing Keil et al.’s PI2 primary DNA and cDNA sequence with JEMI-PI2.

* indicates the location of the intron, + indicates nucleotide base changes relative to Keil et al.’s primary DNA sequence, N represents either an A, T, G, or C nucleotide.

SUPPLEMENTARY FIG. 2. Plasmid maps of (A) pET30ba and (B) pET32a(+)PI2.

Plasmid maps designed in APE program.

SUPPLEMENTAL FIG. 3. Confirming presence of pET30b and ensuring functional Nco1 and Eag1. Single and double digests of pET30b visualized via gel electrophoresis in 1.5% agarose gel with SYBR Safe. 2-log ladder from NEB.

SUPPLEMENTARY FIG. 4. Sequenced PI2 gene in pCR2.1-TOPO-TA.

SUPPLEMENTARY TABLE 1. Strains and plasmids used in this study.

SUPPLEMENTARY TABLE 2. Attempted ligation and transformation of DH5α cells.

REFERENCES


27. **Fermentas.** 2006. Fermentas Restriction Enzymes; Buffer Activities. Waltham, Massachusetts.

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DH5α</td>
<td>Bacterial host for plasmid storage and expression</td>
<td>UBC</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pET30b(+)</td>
<td>Expression vector with kanamycin resistance</td>
<td>Novagen</td>
</tr>
<tr>
<td>pET32a(+)</td>
<td>Expression vector with thioredoxin tag and ampicillin resistance</td>
<td>Novagen</td>
</tr>
<tr>
<td>TOPO-TA</td>
<td>Storage Vector with LacZ multiple cloning site, and ampicillin and kanamycin resistance</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>Vector</td>
<td>Insert</td>
<td>NcoI</td>
</tr>
<tr>
<td>---------</td>
<td>--------</td>
<td>------</td>
</tr>
<tr>
<td>pET30b</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>pET30b</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>pET32a</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>pET32a</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>pET30b</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>pET30b</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>pET30b</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>pET30b</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>pET30b</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>pET30b</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>pET32a</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>pET32a</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>pET30b</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>pET30b</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>pET30b</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>pET30b</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>pUC19 (uncut)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>pET30b (uncut)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>pET32a (uncut)</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>
Participation Report

Arshia Alimohammadi: In charge of the materials and methods, references, general editing, table 1. Hours spent approximately 11 hours.

Jacqueline Siu: In charge of the abstract, parts of the results and discussion, general editing, re-writing of some sections, figure 1, some of figure 3, supplementary figures 1, 2 and 4. Hours spent approximately 14 hours.

Alan Stachowiak: In charge of the introduction, acknowledgements, some of figure 2, general editing, re-writing of some sections, formatting and creation of the manuscript. Hours spent approximately 12 hours.

Emily Fogarty: In charge of parts of the results and discussion, future directions, general editing, section re-writing some of figure 2, 3, supplementary figure 3 and table 2. Hours spent approximately 12 hours.