Volume 6: 1-10

UJEMI+

The Undergraduate Journal of Experimental Microbiology & Immunology (+Peer Reviewed)

Comparative expression of potato proteinase inhibitor type II in an oxidative versus reductive cytosolic environment of *Escherichia coli*

Ravneet Grewal, Wee So Kim, Dini Shi & Haley Tong

Department of Microbiology and Immunology, University of British Columbia, Vancouver, British Columbia, Canada

SUMMARY Potato proteinase inhibitor type II (PI2) is an enzyme primarily expressed in potato tubers that suppresses the activity of serine proteases such as trypsin and chymotrypsin. PI2 is a disulfide-rich protein that requires an oxidative cytosolic redox state to allow for proper folding when expressed in *Escherichia coli.* In this paper, three different strains of *E. coli*, BL21 (DE3), SHuffle and Origami 2, were selected for their distinct cytosolic environments. BL21 (DE3) has a reducing cytosolic environment that is not conducive for disulfide bond formation, whereas Origami 2 and SHuffle have oxidative cytosolic environments that promote disulfide bond formation. Furthermore, SHuffle contains an additional disulfide bond isomerase, DsbC. The objective of this research project was to express and purify PI2 in the three strains of *E. coli* to compare protein folding of the purified product. A vector encoding for maltose-binding protein tagged PI2 was transformed into the three *E. coli* strains and the expression of PI2 was induced. PI2 was subsequently purified on an amylose column and the folded state of the purified protein was probed through a limited proteolysis assay. Results suggest that PI2 was expressed in BL21 and SHuffle, but not in Origami 2. However, purified MBP- PI2 protein was only obtained from SHuffle. Limited proteolysis of the purified protein further suggests that the PI2 purified from SHuffle was in a folded state. Therefore, while BL21 and Origami 2 did not yield distinguishable amounts of MBP-PI2, PI2 can potentially be expressed in a native conformation in SHuffle.

INTRODUCTION

Protease inhibitors are expressed in plant tubers to mediate protective action against damage and stress from insecticidal destruction and viral infections (1). Of these protease inhibitors, potato proteinase inhibitor type II (PI2) is one type which operates by suppressing the activity of common serine proteases, trypsin and chymotrypsin (1). Among its other roles, PI2 has also been found to induce appetite suppression and exhibits anti-metastatic effects on human colorectal cancer (2, 3). Thus, researchers have been interested in high-throughput expression of PI2 in *Escherichia coli* to facilitate further investigations on potential therapeutic applications (4-8).

PI2 is composed of two domains that both contribute to its protease inhibitory activity (1). The first active site inhibits either chymotrypsin or trypsin, as determined by the variant of PI2, but not both (1). Conversely, the second active site has broader activity, with the capacity to inhibit both chymotrypsin and trypsin (1). The specificity of PI2 for serine proteases is defined by the motif, CxxxC, which is conserved in all PI2 variants (1).

Published Online: September 2020

Citation: Ravneet Grewal, Wee So Kim, Dini Shi & Haley Tong. 2020. Comparative expression of potato proteinase inhibitor type II in an oxidative versus reductive cytosolic environment of *Escherichia coli*. UJEMI+ 5:1-10

Editor: Daniela Morales and Julia Huggins, University of British Columbia

Address correspondence to: https://jemi.microbiology.ubc.ca/

Altogether, PI2 contains eight disulfide bridges contributing to its structure, which are formed through the oxidation of thiol groups in cysteine residue pairs (9). Disulfide bridges are often highly involved in maintaining protein stability (9). Therefore, if the pairing of these disulfide bonds is prevented by the cytosolic redox state of the host cell, it may result in the formation of random protein aggregates rather than the properly folded protein. However, in many strains of *E. coli* commonly used for protein production, such as BL21 (DE3), these disulfide bridges may not be formed properly given the reducing cytosolic environment (10, 11). Due to the efficiency of *E. coli* as a host for recombinant protein production, certain specialized strains for disulfide-containing protein expression have been developed in response to this problem (12). One such strain is the *E. coli* K-12 variant Origami 2 (DE3), which has been engineered to have an oxidizing cytosolic compartment (11). A study by Ang *et al.* attempted to express PI2 in the Origami 2 strain but observed a lack of PI2 in the cell lysates (8). As an alternative to Origami 2, New England BioLabs developed SHuffle® Express Competent *E. coli* (C3028). Unlike Origami 2, the SHuffle strain expresses a prokaryotic disulfide bond isomerase, DsbC, which assists in the refolding of mis-oxidized proteins, in addition to having an oxidative cytosolic environment (13, 14). Experiments involving other proteins with disulfide bonds have demonstrated higher expression levels in SHuffle cells compared to Origami 2 cells (14). Therefore, SHuffle cells may provide a more suitable environment for the expression of conformationally active PI2.

The aim of this paper was to express PI2 in *E. coli* strains BL21 (DE3), Origami 2 and SHuffle (C3028) and to subsequently investigate the proper folding of the purified protein products by performing limited proteolysis using Proteinase K. We hypothesized that MBP-PI2 expressed in the oxidative cytosol of SHuffle cells will have a properly folded conformation, as the disulfide bonds in PI2 will be able to pair correctly with assistance from both DsbC and the oxidative environment. However, PI2 expressed in the reductive cytosol of BL21 cells will not be properly folded due to the unfavourable redox environment. Moreover, the absence of DsbC in both Origami 2 and BL21 will result in less properly folded PI2. Through induction experiments in all three strains, PI2 expression was only detected in BL21 and SHuffle. PI2 was subsequently purified from only SHuffle. Results of a limited proteolysis assay on the purified protein revealed that MBP-PI2 may be in a uniformly folded state. This data supports the notion that compared to other strains investigated, SHuffle is indeed a more suitable strain for future production of MBP-PI2 in *E. coli*.

METHODS AND MATERIALS

Bacterial strains and plasmids. *E. coli* strains BL21 (DE3) and Origami 2 were cultured from frozen glycerol stocks provided by the University of British Columbia (UBC), Department of Microbiology and Immunology. Competent *E. coli* strain SHuffle (C3028) was provided by the Straus Lab at UBC. All cultures were streaked onto LB agar plates and subsequently grown in LB broth. Origami 2 and BL21 (DE3) were grown at 37°C and SHuffle was grown at 30°C. A summary of the strains used was compiled in Table S1.

Plasmid pMAL-c2X was provided by the Eltis Lab at UBC, while pMAL-c2X-LLMZ16 (hereby referred to as pMAL-c2x-PI2) was provided by Lapointe *et al.* (Table S1). Both plasmids were maintained in *E. coli* strain DH5α and were streaked onto LB agar containing ampicillin (100 μg/mL) and grown at 37°C. The pMAL-c2x and pMAL-c2x-PI2 plasmids were isolated from DH5α using the BioBasic EZ-10 Spin Column Plasmid DNA Miniprep Kit, following manufacturer protocol. Isolated plasmids were sent to GeneWiz for Sanger sequencing using the universal M13 primers. The sequences were then aligned in NCBI BLAST against the pMAL-c2x-LLMZ16 sequence provided by Lapointe *et al*. A vector map was created, and the alignment was conducted using Geneious Prime (Fig. S1).

Preparation of competent *E. coli*. BL21 (DE3), SHuffle and Origami 2 *E. coli* strains were made competent following the procedure described by Chang *et al*. Cells were grown to an OD_{600} of 0.4 and centrifuged at 4000 rpm at 4°C for 10 minutes. The resulting cell pellet was resuspended in ice-cold 0.1 M $CaCl_2$ and set on ice for 30 minutes. The centrifugation process was repeated, and the pellets were resuspended in 5 mL of 0.1 M $CaCl_2$ with 15% glycerol. Samples were aliquoted and stored at -80°C for future use in transformations.

Transformation of pMAL-c2x and pMAL-c2x-PI2 into *E. coli*. The pMAL-c2x and pMAL-c2x-PI2 plasmids were transformed into competent *E. coli* as per the method described by Chang *et al*. Purified pMAL-c2x or pMAL-c2x-PI2 was added to 100 µL of each competent *E. coli* strain and placed on ice for 30 minutes. Samples were subsequently heat shocked for 45 seconds at 42°C and placed on ice for 2 minutes. LB broth was then added for a final dilution of 1/10. BL21 (DE3) and Origami 2 cells were incubated at 37°C, while SHuffle cells were incubated at 30°C in a shaking incubator for 1 hour. All transformants were plated on LB agar containing ampicillin (100 µg/mL). A single colony was picked from each transformed strain and grown overnight in LB Broth containing ampicillin (100 µg/mL). Each overnight culture was mixed in a 1:1 ratio with 50% glycerol stock and stored at -80°C.

Growth curves. Growth measurements were made over a 20-hour period to confirm that there were no significant growth defects in SHuffle and BL21 (DE3) *E. coli* strains, and that the two vectors did not result in significant differences in growth (Fig. S2). Overnight cultures of strains were diluted to an OD_{600} of 0.2 and plated in triplicate into a 96-well plate. The plate was incubated at 30°C and set to orbital shake continuously at a frequency of 282 CPM, while the BioTek Epoch 2 Microplate Reader read the OD_{600} every 30 minutes for 20 hours.

Induction of PI2 expression. Single colonies of transformed BL21 (DE3) and SHuffle were respectively inoculated in 5 mL of LB broth containing ampicillin (100 µg/mL). 1 mL of each overnight culture was inoculated into 100 mL of LB broth containing ampicillin (100 µg/mL) and 2% glucose to suppress the expression of maltose genes and the lac operon. Origami 2 was also induced in LB broth containing ampicillin (100 µg/mL) and no glucose supplement. All cultures were grown at 30°C until the sample reached an OD_{600} of 0.6. The cultures were then induced with 1 mM of IPTG and incubated overnight at 30°C, shaking at 200 rpm. Following induction, aliquots of each induced culture were centrifuged at 16,000 x g for 10 minutes. The pellets were resuspended in lysis buffer (20 mM Tris (pH 7.4), 200 mM NaCl, 10% glycerol, 1 mM EDTA, 1 mM DTT) and incubated on a room temperature shaker at 250 rpm for 10 minutes. Samples were run on SDS-PAGE as described below. The remaining cells were pelleted by centrifugation at 2000 x g for 15 minutes and the resulting pellets were frozen at -80°C.

SDS-PAGE. Samples were mixed with BioRad 2X Laemmli buffer (Cat no. 161-0737) containing 2-mercaptoethanol and heated at 95°C for 5 minutes. The proteins present in each cell lysate were resolved by SDS-PAGE on a 4-20% Bio-Rad Mini-PROTEAN® TGX Stain-Free™ Precast Gel (Cat no. #456-8084) with a Tris-Glycine buffer system. Protein bands were compared to BioRad Precision Plus Protein™ Unstained Standards (Cat no. #1610363) and 2 mg/ml GBiosciences BSA Standard (Cat no. 224B-B). Putative MBP-PI2 bands intensities were normalized to the total intensity of the samples using ImageJ (16).

MBP-PI2 Purification from SHuffle and BL21 (DE3). Frozen cell pellets of induced BL21 (DE3) and SHuffle transformants were resuspended in 2 mL of lysis buffer containing VWR Protease Inhibitor Cocktail (Cat no. 97063-970). Cells were homogenized using the MP FastPrep-24 homogenizer at 6 m/s for 6 seconds. The silica beads and cell debris were pelleted and removed through centrifugation at 15,000 x g for 10 min. The resulting supernatant, containing the soluble protein fraction, was collected for subsequent purification.

A 1 mL amylose resin (New England BioLabs, Cat no. E8021S) gravity flow column was used to carry out purification. Five column volumes (CV) of column buffer (20 mM Tris (pH 7.4), 200 mM NaCl, 10% glycerol, 1 mM EDTA) were used to equilibrate the column. The soluble protein fraction from each pellet was incubated with amylose resin for 1 hour at 4°C. The column was then washed with 10 CV of Column Buffer, before the MBP-PI2 was eluted with 5 CV of elution buffer (20 mM Tris (pH 7.4), 200 mM NaCl, 10% glycerol, 1 mM EDTA, 10 mM maltose). Fractions were analyzed by SDS-PAGE, as previously discussed. BioRad Precision Plus Protein™ Unstained Standards (Cat no. #1610363) and 2 mg/ml GBiosciences BSA Standard (Cat no. 224B-B) were included as molecular weight ladders. Select elution

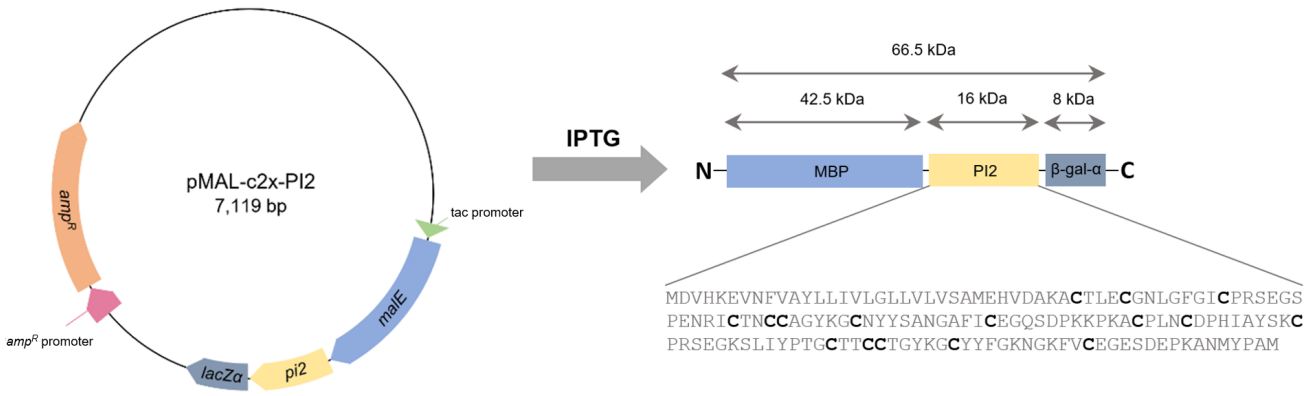

FIG. 1 Expression vector construct of pMAL-c2x-PI2. The pMAL-c2x-PI2 vector was cloned by Lapointe *et al*. through the insertion of a *pi2* sequence variant with spliced out introns and optimized GC content for expression in *E. coli*. Inducible with IPTG, the resulting fusion protein is expected to be 66.5 kDa, consisting of a PI2 insert (16 kDa) flanked by a maltose-binding protein tag (42.5 kDa) at the N-terminus end and a ß-gal-α protein fragment at the C-terminus end. A total of 16 cysteine residues were found in the PI2 protein sequence, all expected to be forming disulphide bridges.

fractions from induced SHuffle containing pMAL-c2x-PI2 were pooled together and the protein was quantified using the NanoDrop™ 2000c Spectrophotometer (Thermo Scientific) using a molar extinction coefficient of 82250 M^{-1} cm^{-1}.

Limited Proteolysis of Purified MBP-PI2 Protein using Proteinase K. Proteinase K (1 µg/mL) was added to purified MBP-PI2 protein (48 µg/mL) at a 1:1 volume and incubated at 40°C to activate the reaction. Aliquots of the reaction were taken over several time points. These aliquots were then resolved by SDS-PAGE, as previously described, to visualize the gradual degradation of MBP-PI2. BioRad Precision Plus Protein™ Unstained Standards (Cat no. #1610363) and 2 mg/ml GBiosciences BSA Standard (Cat no. 224B-B) were included as molecular weight ladders.

RESULTS

Sequencing of the pMAL-c2x-PI2 plasmid construct. Previously, Lapointe *et al.* created the pMAL-c2X-PI2 plasmid construct containing the *pi2* gene variant, synthesized by Fogarty *et al.*, that was optimized for expression in *E. coli* by removing introns in the *pi2* reading frame. The construct was N-terminally tagged with a maltose-binding protein (MBP) tag in an attempt to improve the overall solubility of the protein and to facilitate subsequent purification steps (7). In order to confirm that the pMAL-c2x-PI2 plasmid cloned by Lapointe *et al*. contained the *pi2* gene insert, the plasmid was sequenced with the M13 forward primer. The BLAST alignment tool was used to align the Sanger sequencing results to the optimized *pi2* gene sequence (6). The results of this alignment revealed that the *pi2* gene was absent in the pMAL-c2x empty vector plasmid, as expected. However, the same gene was identified within the pMAL-c2x-PI2 sequence and shown to be aligning in the correct orientation. A vector map of pMAL-c2x-PI2 was subsequently generated using Geneious Prime (Fig. S1). The *pi2* gene was shown to be inserted in between the *malE* gene and the *lacZα* gene, encoding for an N-terminal MBP-tag and a C-terminal ß-gal-α protein fragment, respectively (Fig. 1). These genes are all downstream of a *tac* promoter, which is inducible by IPTG (Fig. 1). An ampicillin resistance marker was also present in a separate region of the plasmid (Fig. 1). ExPASy translation of the *pi2* gene insert confirmed the presence of 16 cysteine residues within the amino acid sequence, which is consistent with literature (Fig. 1) (17). Overall, the *pi2* gene was found to be inserted correctly in the pMAL-c2x-PI2 plasmid for MBP-PI2 expression.

MBP-PI2 was expressed in BL21 (DE3) and SHuffle, but not Origami 2. *E. coli* strains BL21 (DE3), SHuffle and Origami 2 transformed with pMAL-c2x-PI2 were induced with IPTG to assess MBP-PI2 expression levels in each host strain. Negative controls consisted of

uninduced PI2 expression vector, along with induced empty vector transformed cells. SDS-PAGE analysis of BL21 (DE3) empty vector crude cell lysate showed overexpression of a protein at approximately 50 kDa (Fig. 2). This may represent MBP-β-gal-α, which has a predicted size of 50.5 kDa (18). Conversely, this expressed protein is not as distinctly overexpressed in the SHuffle or Origami 2 empty vector lysates (Fig. 2). SDS-PAGE analysis also revealed a distinct protein band at the expected molecular weight of MBP-PI2 (66.5 kDa) in the induced BL21 (DE3) PI2 vector lysate, which is consistent with the results previously obtained by Lapointe *et al.* and Ang *et al.* (7, 8). Moreover, this MBP-PI2 band was also apparent in the SHuffle sample under the same test conditions and was more concentrated compared to BL21 (DE3) as quantified by ImageJ (Fig. 2) (16). This band was undetected in the uninduced PI2 vector and induced empty vector lysates for both BL21 and SHuffle (Fig. 2). Our results suggest that MBP-PI2 expression is inducible in BL21 (DE3) and SHuffle, with relatively higher levels of expression in SHuffle.

Figure 2 reveals no distinct 66.5 kDa MBP-PI2 band in the induced PI2 vector-transformed Origami 2 lysate, as was expected based on previous experiments by Ang *et al*. Interestingly, modifying media conditions by removing the 2% glucose supplement resulted in an increase in protein expression in Origami 2 (Fig. S3). Most notably, expression of an unknown protein of approximately 50 kDa was observed in induced Origami 2 lysates containing the PI2 and empty vector, while expression of a protein at approximately 66 kDa was observed in the uninduced Origami PI2 vector lysate (Fig. S3). However, as the 66.5 kDa MBP-PI2 band remained absent in the induced Origami 2 PI2 vector lysate grown with and without glucose supplement, these results suggest that Origami 2 may not be a suitable host for expressing MBP-PI2 (Fig. 2, S3).

Purification of MBP-PI2 from BL21 (DE3) and SHuffle. Since expression of MBP-PI2 was only observed in BL21 (DE3) and SHuffle cells, MBP-PI2 was purified from both strains for further characterization. The protein content of each elution fraction was visualized through SDS-PAGE analysis. As expected, the elution fractions from the empty vector cultures did not contain a 66.5 kDa MBP-PI2 band (Fig. 3). However, the purification of a 50.5 kDa protein, likely MBP-β-gal-α, was observed in the BL21 empty vector elution fractions (Fig. 3). Conversely, this protein was absent from the SHuffle empty vector elution fractions, which reflects the trends observed in the whole-cell lysate (Fig. 2, 3). The SHuffle empty vector purification yielded a band that matches the molecular weight of unmodified MBP (42.5 kDa) (Fig. 3). Surprisingly, despite the 66.5 kDa band observed in the BL21 (DE3) PI2 vector lysate in the induction gel and the previous successful MBP-PI2 purification conducted by Lapointe *et al.*, we were unable to purify MBP-PI2 from BL21 (DE3) (Fig 2, 3). Instead, a 42.5 kDa protein, likely to be MBP, appeared to have been isolated (Fig. 3). On the other hand, the 66.5 kDa MBP-PI2 band was distinctly present in the elution fractions from the SHuffle PI2

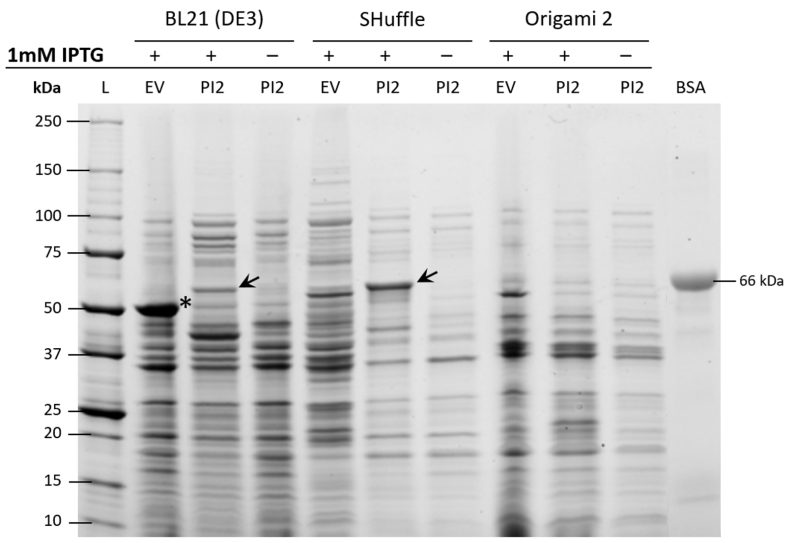

FIG. 2 MBP-PI2 was expressed in BL21 (DE3) and SHuffle, but not Origami 2. BL21 (DE3), SHuffle and Origami 2 cells transformed with pMAL-c2X (EV) or pMAL-c2X-PI2 (PI2) were induced overnight with 1 mM IPTG at 30°C. The resulting cell lysates were analyzed by SDS-PAGE. A reference ladder (L) and bovine serum albumin (BSA) served as molecular weight standards. Bands at the expected molecular weight of MBP-PI2 (66.5 kDa) were seen in both the induced PI2-transformed BL21 (DE3) and SHuffle lysates, as indicated by the arrows. However, this band was absent in the induced PI2-transformed Origami 2 lysate. The MBP-PI2 band had a higher intensity in SHuffle compared to BL21 (DE3), suggesting higher MBP-PI2 expression in SHuffle. All induced EV-transformed and uninduced PI2-transformed cells lacked the MBP-PI2 band. Furthermore, the asterisk may represent MBP-β-gal-α (50.5 kDa) that is expressed by the EV-transformed BL21 that does not appear in EV-transformed SHuffle.

transformant (Fig. 3). However, smaller molecular weight protein bands were also apparent in the fractions (Fig. 3). The first elution fraction of PI2 expressed in SHuffle contained too many additional proteins and was excluded in the final pooled purified product. Nonetheless, MBP-PI2 was successfully purified from SHuffle.

Limited Proteolysis Assay on MBP-PI2. Finally, to probe further into the folding of the MBP-PI2 protein, elution fractions 2-4 were pooled for a limited proteolysis assay conducted using Proteinase K. SDS-PAGE gel analysis showed a clear 66.5 kDa protein band in the purified MBP-PI2 product, prior to the addition of Proteinase K (Fig. 4). The same 42 kDa off-target protein from the purification gel was present in the negative control lane of the limited proteolysis gel, which was expected, as no further purification steps were undertaken (Fig. 3, 4). A distinct banding pattern for MBP-PI2 was observed after gradual degradation by Proteinase K (Fig. 4). Since Proteinase K is expected to cleave more exposed areas of the protein, a uniformly folded protein sample should result in a consistent banding pattern. On the other hand, randomly aggregated proteins would have been cleaved at arbitrarily exposed sites, and the resulting random peptide fragments would not have been at high enough concentration to show a distinct banding pattern. Specifically, the 66.5 kDa MBP-PI2 band became fainter over time, while a band at approximately 42.5 kDa, the expected size of MBP, became more intense over time (Fig. 4B). This may suggest that the protein has folded such that the linker between the MBP and PI2 domain is exposed and being cleaved throughout this time course, resulting in a decrease in MBP-PI2 and an increase in MBP. Proteins below 20 kDa, where the predicted 16 kDa PI2 would have been expected, were not resolved (Fig. 4A). The overall consistency of the band pattern observed across all exposures to Proteinase K suggest that the SHuffle-purified, MBP-tagged protein is exhibiting proper protein folding.

DISCUSSION

Protease inhibitors such as PI2 are found in many plants and are known for their protective action against infection and environmental stress (1). PI2 is of interest to the scientific community as its functionality extends to human therapeutic applications such as preventing metastasis of colorectal cancer (3). However, the expression of functional PI2 in prokaryotic hosts has proven to be challenging, as the eight disulfide bridges in the three-dimensional folded state of the protein may not pair correctly in the reductive state of the cytosol (9, 12). Therefore, the expression of PI2 in hosts with an oxidative cytosolic compartment has been investigated in order to promote correct pairing of disulfide bonds and overall proper folding of the protein.

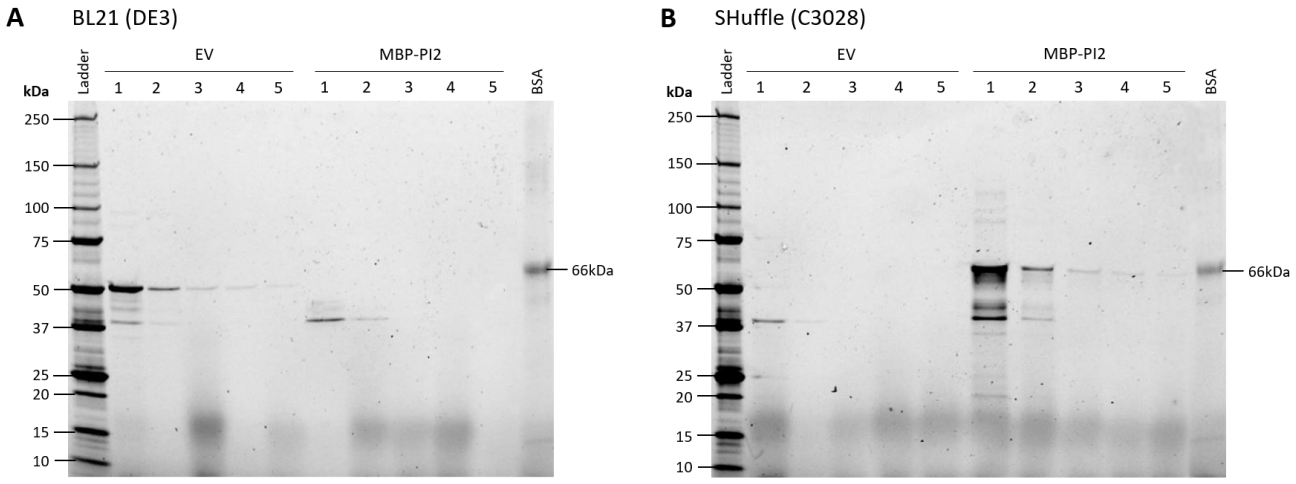

FIG. 3 MBP-PI2 was purified from SHuffle cells. Soluble protein fractions from induced EV and PI2-transformed BL21 (DE3) (A) and SHuffle (B) were purified using amylose affinity chromatography. The elution fractions were analyzed by SDS-PAGE. A reference ladder and bovine serum albumin (BSA) served as molecular weight standards. MBP-PI2 (66.5 kDa) was observed in all elution fractions from PI2-transformed SHuffle, but absent in elution fractions from PI2-transformed BL21 (DE3). Some lower molecular weight proteins also appeared to be co-purified with MBP-PI2 from SHuffle. MBP-PI2 was absent in elution fractions from all EV-transformed cells. However, MBP-β-gal-α (50.5 kDa) was purified from EV-transformed BL21 (DE3).

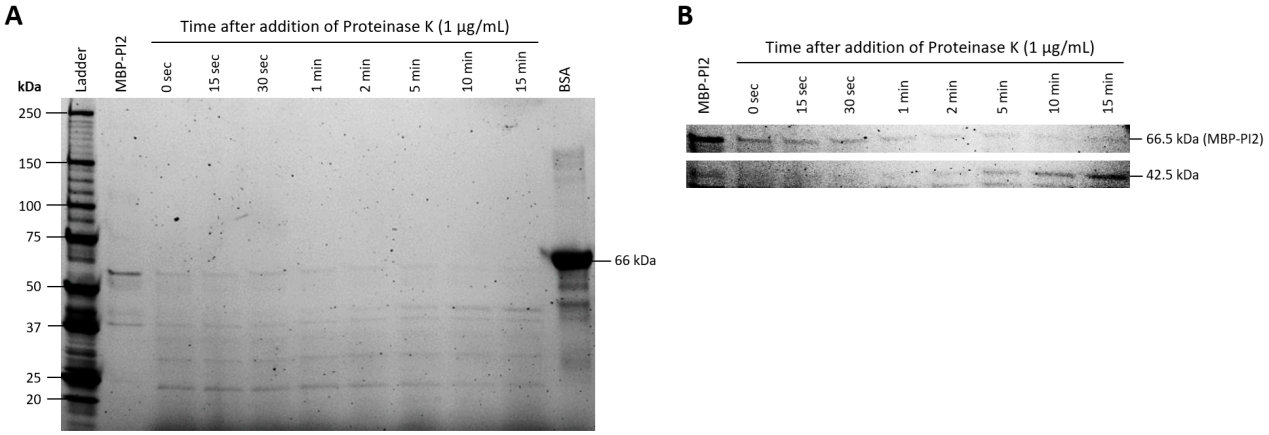

FIG. 4 Limited proteolysis assay of purified MBP-PI2 from SHuffle. Purified MBP-PI2 from SHuffle was digested with 1 μg/ml Proteinase K and samples of the reaction were taken at various time points ranging from 0 to 15 minutes. Samples were analyzed by SDS-PAGE. A reference ladder and bovine serum albumin (BSA) served as molecular weight standards. The full gel showed a distinct banding pattern forming throughout the reaction (A). Enhancing contrast of the gel image shows a decrease in intensity of a 66.5 kDa MBP-PI2 protein band over time, corresponding with an increase in intensity of a 42 kDa MBP band (B).

Previously, Ang *et al.* sought to compare the expression of MBP-PI2 in the reductive cytosol of BL21 (DE3) to the oxidative cytosol of Origami 2. However, no expression of MBP-PI2 was detected in Origami 2, contrary to their hypothesis (8). In this paper, the expression and isolation of MBP-PI2 from another host with an oxidative cytosol, SHuffle, was described and compared to the previous strains. Results revealed that PI2 was expressed in SHuffle and BL21 (DE3), but not in Origami 2 (Fig. 2). This supports conclusions made by Ang *et al*. However, despite using the optimal conditions described by Lapointe *et al.* and Ang *et al.*, MBP-PI2 expression in BL21 (DE3) with expression vector appeared to be fainter than in SHuffle (Fig. 2). The presence of the disulfide bond isomerase, DsbC, in SHuffle may have played a role in further stabilizing and correctly pairing the disulfide bridges during protein folding. Therefore, higher concentrations of PI2 may have been seen in SHuffle compared to both BL21 (DE3) and Origami 2 due to more misfolded PI2 being degraded in the other two strains before detection.

The lack of MBP-PI2 expression in Origami 2 observed both in this study and by Ang *et al.* may also be a result of changes in resource allocation in the cell upon addition of the PI2 plasmid construct. There are finite resources in the host cell available for biological processes such as the transcription and translation of proteins. These resources must be allocated precisely to optimize cell fitness (21). In Origami 2, resources may simply be allocated in a way such that MBP-PI2 cannot expressed while simultaneously accommodating essential cell functions. However, when Origami 2 containing the PI2 vector or the empty vector was induced without glucose supplement in the media, there was high expression of an unknown protein at approximately 50 kDa that was not found in Origami 2 induced in glucose-supplemented media (Fig. S3). The addition of glucose in the induction media was used to repress both the expression of maltose genes as well as the lac operon in the chromosomes of host cells (8, 18). Suppression of the maltose operon and its gene targets was necessary, as one of the proteins encoded is amylase, which may have degraded the amylose on the affinity column in downstream purification steps (19). Since amylase has a molecular weight of 51-54 kDa, it may be amylase protein that was observed in the gel (Fig. S3) (20). While glucose-free media may not make Origami 2 more likely to express MBP-PI2, the result of this experiment showed that changing the media composition directly shifted the resource allocation in the cell to result in the expression of a different set of proteins. Therefore, perhaps a different media composition may be explored to allow for the expression of PI2 in Origami 2 strain.

Following the induction experiments, purification of MBP-PI2 was obtained in SHuffle but not in BL21 (DE3) (Fig. 3). This is inconsistent with results obtained by Lapointe *et al.*, who were able to purify MBP-PI2 in BL21 (DE3) (7). However, given the lower intensity the putative MBP-PI2 band induced in BL21 (DE3) (Fig. 2), MBP-PI2 may have been lost during the purification process, as the affinity chromatography method used was the only purification strategy employed and no further optimization steps were taken. It is possible that the PI2

fragment of the MBP-PI2 protein may have been unstable in the reductive cytosol due to problems with disulfide bond formation. This may have resulted in the PI2 segment of the improperly folded protein blocking the MBP binding site, decreasing the affinity of the MBP tag to the amylose column (22).

Off-target proteins or molecules were also found to be eluting along with MBP-PI2, as lower molecular weight bands were seen in the gel (Fig. 3). Since the column binds the MBP tag, the lower eluting bands are likely truncated proteins. The translation of MBP-PI2 may not have been completed or may have been disrupted. Since the MBP tag is translated first, it is reasonable to assume that the bands visualized just above 37 kDa, at around 42.5 kDa, may be the MBP tag only without PI2 (Fig. 3). Furthermore, it is possible that the linker between the MBP tag and PI2 protein is exposed and susceptible to cleavage by host cell proteases. In addition, the protease inhibitor cocktail was only added in the cell lysis stage, and not during subsequent resin incubation or elution phases. If MBP-PI2 was not very stable, then it may have been susceptible to protease degradation during those steps. However, the decision to remove protease inhibitors in latter steps was due to PI2 itself being a protease inhibitor, so the addition of the protease inhibitor cocktail would have possibly disrupted the downstream limited proteolysis assay. Nonetheless, this implies that contaminating proteases may have been introduced into the sample at some point during the purification process, which may explain the smaller bands that may have derived from cleavage. These contaminant proteins may have been removed using size exclusion chromatography, but as their presence was minimal and MBP-PI2 was still the most dominant protein, the eluted fractions from affinity chromatography were subsequently used for proteolysis analysis.

Proteinase K-treated MBP-PI2 displayed a consistent cleavage pattern (Fig. 4). The intensity of the putative MBP-PI2 band at 66.5 kDa decreased in a time-dependent manner with Proteinase K, suggesting that MBP-PI2 was subjected to degradation (Fig. 4). Furthermore, the change in intensity of a 42.5 kDa band was inversely related to the 66.5 kDa band (Fig. 4). This may be reflecting an increase in concentration of the MBP tag, which may indicate that the linker between MBP and PI2 is exposed for cleavage by proteinase K and that MBP and PI2 fold into distinct domains. However, the gel did not resolve below 20 kDa. Therefore, while an increase in intensity of the band at 42.5 kDa may potentially suggest cleavage of MBP-PI2 into the MBP tag and PI2 protein, a complementary increase in intensity of a 16 kDa PI2 band cannot be confirmed. Use of an SDS-PAGE gel with higher acrylamide concentration may help resolve these smaller proteins fragments in future studies. Nonetheless, the presence of a consistent cleavage pattern and absence of a protein smear may suggest that the purified MBP-PI2 product is homogenous as Proteinase K cuts at similar sites each time. Therefore, the purified MBP-PI2 is likely in a uniform folded conformation and not a random aggregate of protein.

The lack of misfolded MBP-PI2 may be a result of the oxidative environment provided by SHuffle. It has previously been found that oxidative environments are necessary for the formation of disulfide bonds, which allow for subsequent folding of the protein into its proper conformation (9, 12). Expression of disulfide-containing proteins in reductive environments have been found to result in the mispairing or partial oxidation of the thiol groups of cysteine residue pairs, which lead to the formation of inclusion bodies consisting of protein aggregates (9, 12). As such, it is theorized that the combination of the oxidative environment of SHuffle and the presence of DsbC isomerase facilitated the formation of the correct disulfide bonds in MBP-PI2 (13). This would have in turn increased the overall stability of MBP-PI2 in a stable folded state instead of a random aggregate, which was observed through limited proteolysis of MBP-PI2 (Fig. 4). A parallel study with MBP-PI2 purified from a DsbC isomerase-expressing *E. coli* strain with a reductive cytosol would be required to further support this theory.

This study demonstrates that SHuffle (C3028), a strain of *E. coli* that has been engineered to have an oxidative cytosol with a DsbC isomerase, can produce MBP-PI2 that may be uniformly folded based on a preliminary limited proteolysis experiment. The presence of disulfide bond isomerase, DsbC, which is absent in other oxidative strains like Origami 2, may play an important role in optimizing conditions for folding and pairing of cysteine residues in the PI2 protein. In this paper, the structural information of MBP-PI2 expressed in a reductive cytosol was unable to be compared, as no MBP-PI2 was purified from BL21 (DE3). While the functionality and correct folding of MBP-PI2 cannot be verified in the current study, this paper

paves the way for further investigation to be done on the functional activity of PI2 expressed in SHuffle and provides a solid foundation for the production of PI2 in an *E. coli* system.

Many cellular processes are regulated or are influenced by serine proteases. Production of active PI2 through an efficient, low cost host such as *E. coli* may facilitate further studies on possible therapeutics in anti-cancer and anti-viral applications (1-3). Similar serine protease inhibitors are already being used to treat pathogenic infections and blood coagulation disorders (23). Thus, the development of novel drug candidates derived from PI2, to selectively inhibit mutant serine proteases is an important area of study (23). PI2 may be implicated in many possible applications and further characterization of its activity is needed to elucidate its potential utility.

Future Directions Future studies may consider optimizing the induction and purification process for PI2 expression in SHuffle cells. Parameters such as induction time and temperature may be adjusted to yield maximal PI2 expression. Further purification steps, such as size exclusion chromatography, may improve the purity of the extracted protein. Similar adjustments on the induction and purification of MBP-PI2 from BL21 (DE3) would be required to compare the structure and activity of PI2 expressed in a reductive environment. Western blot analysis using anti-PI2 antibody may also be used to confirm the expression of MBP-PI2 in SHuffle and BL21 (DE3) strains. This would also allow for more reliable results for any downstream structural and functionality analysis.

For further confirmation that MBP-PI2 is in a folded state, the protease inhibition activity of PI2 can be assessed through a functionality test, such as a trypsin inhibition assay. Given that cysteine residues are involved in the active site of PI2, it is hypothesized that the correct formation of disulfide bonds is required for the site to fold into an active conformation (1). It is also important to consider the MBP affinity tag linked to the N-terminus of PI2 (7). Since the MBP tag (42.5 kDa) is much larger compared to PI2 (kDa), it may impede protein-protein interactions required for PI2 activity (1, 7). Therefore, it may also be of interest to cleave off the MBP tag in future studies using the Factor Xa protease. Factor Xa has the capacity to remove MBP tags from recombinant proteins following purification (18). Removal of the tag would potentially allow for a more accurate investigation into the proper folding and functionality PI2 in different *E. coli* hosts. This would subsequently support further research into therapeutic applications of PI2.

ACKNOWLEDGEMENTS

We would like to express our gratitude to Straus lab for providing the SHuffle C3028 strain, the Eltis Lab for providing the pMAL-c2x plasmid and Lapointe *et al.* for providing the pMAL-c2x-LLMZ16 plasmid. We would also like to thank the UBC Department of Microbiology and Immunology for funding this project. Special thanks to Dr. David Oliver, Evelyn Sun and Reynold Calderon for their support, guidance and willingness to discuss our project. Without them, this project would not have been possible. We would also like to thank two anonymous reviewers for constructive feedback on this manuscript.

REFERENCES

1. **Beekwilder J, Schipper B, Bakker P, Bosch D, Jongsma M.** 2000. Characterization of potato proteinase inhibitor II reactive site mutants. Eur J Biochem **267:**1975-1984.
2. **Komarnytsky S, Cook A, Raskin I.** 2011. Potato protease inhibitors inhibit food intake and increase circulating cholecystokinin levels by a trypsin-dependent mechanism. Int J Obesity **35:**236–243.
3. **Li PG, Mu TH, Deng L.** 2013. Anticancer effects of sweet potato protein on human colorectal cancer cells. World J Gastroenterol **19:**3300-3308.
4. **Geum L, Huber R, Leung N, Lowe M.** 2015. Construction of Recombinant Expression Vectors to Study the Effect of Thioredoxin on Heterologous Protein Solubility. JEMI **10:**1-5.
5. **Przeworski C, Pham D, Wang, I, Murillo J.** 2015. Attempted Construction of Recombinant Vectors Designed to Study the Solubility of Overexpressed Proteinase Inhibitor 2 when Coexpressed with Thioredoxin. JEMI **19:**1-6.
6. **Fogarty E, Alimohammadi A, Siu J, Stachowiak A.** 2016. Synthesis, Cloning, and Sequencing of a Codon Optimized Variant of Proteinase Inhibitor II Designed for Expression in *Escherichia coli.* JEMI **20:**100-105.

7. **Lapointe HR, Li S, Mortazavi S, Zeng J.** 2016. Expression and purification of a potato type II proteinase inhibitor in *Escherichia coli* strain BL21(DE3). JEMI+ **2:**34-40.
8. **Ang I, Atte A, Halim B, Jassal J.** 2017. Effect of Temperature, Inducer Concentration, and *Escherichia coli* Cytosolic Redox State on MBP-PI2 Expression. JEMI **21:**11-14.
9. **Li XQ, Zhang T, Donnelly D.** 2011. Selective Loss of Cysteine Residues and Disulphide Bonds in Potato Proteinase Inhibitor II Family. PLoS ONE **6:**e18615.
10. **Yoon, SH, Jeong, H, Kwon, S, Kim, JF.** 2009. Genomics, Biological Features, and Biotechnological Applications of *Escherichia coli* B: "Is B for better?!", p. 1-17. *In* S. Y. Lee (ed.), Systems Biology and Biotechnology of *Escherichia coli*. Springer Netherlands, Dordrecht.
11. **Xiong S, Wang YF, Ren XR, Li B, Zhang MY, Luo Y, Zhang L, Xie QL, Su KY.** 2005. Solubility of disulfide-bonded proteins in the cytoplasm of *Escherichia coli* and its "oxidizing" mutant. World J Gastroenterol **11:**1077-1082.
12. **Berkmen, M.** 2012. Production of disulfide-bonded proteins in *Escherichia coli.* Protein Expr Purif **82:**240-251.
13. **Lobstein J, Emrich CA, Jeans C, Faulkner M, Riggs P, Berkmen M.** 2012. SHuffle, a novel *Escherichia coli* protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. Microb Cell Fact **11:**56.
14. **New England Biolabs.** 2019. SHuffle® Express Competent *E. coli*. New England Biolabs, Ipswich, MA.
15. **Chang AY, Chau VWY, Landas JA, Pang Y.** 2017. Preparation of calcium competent *Escherichia coli* and heat-shock transformation. JEMI methods **1:**22-25.
16. **Schneider CA, Rasband WS, Eliceiri, KW.** 2012. NIH Image to ImageJ: 25 years of image analysis. Nat methods **9**: 671-675.
17. **Keil M, Sanchez-Serrano J, Schell J, Willmitzer L.** 1986. Primary structure of a proteinase inhibitor II gene from potato (Solanum tuberosum). Nucleic Acids Res **14**:5641–5650.
18. **New England BioLabs.** 2003. pMAL™ Protein Fusion and Purification System (Version 5.1). New England BioLabs, Ipswich, MA.
19. **Griffiths AJF, Miller JH, Suzuki DT, Lewontin RC, Gelbart WM.** 2000. Introduction to Genetic Analysis. W. H. Freeman.
20. **Cozzone P, Paséro L, Marchis-Mouren G.** 1970. Characterization of porcine pancreatic isoamylases: separation and amino acid composition. Biochim Biophys Acta **200**:590–593.
21. **Borkowski, O, Ceroni, F, Stan, G, Ellis, T.** 2016. Overloaded and stressed: whole-cell considerations for bacterial synthetic biology. Curr Opin Microbiol **33:**123-130.
22. **Costa, S, Almeida, A, Castro, A, Domingues, L.** 2014. Fusion tags for protein solubility, purification and immunogenicity in *Escherichia coli*: the novel Fh8 system. Front Microbiol. **5:**63.
23. **Harish BS, Uppuluri KB.** 2018. Microbial serine protease inhibitors and their therapeutic applications. Int J Biol Macromol **107**:1373–1387