Unlocking the Potential of Extensin Signal Peptide and Elastin-like polypeptide Tag fused to Shigella dysenteriae’s IpaDSTxB to Improve Protein Expression and Purification in Nicotiana tabacum and Medicago sativa

AmirMohammad Soleimani
A.m.soleimani12@gmail.com

Houshang Alizadeh
University of Tehran

Research Article

Keywords: Apoplastic and Symplastic protein, Elastin like polypeptide, Extensin signal peptide, Protein extraction and Purification, Shigellosis

Posted Date: April 17th, 2024

DOI: https://doi.org/10.21203/rs.3.rs-4255517/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Protein Expression and Purification on June 1st, 2024. See the published version at https://doi.org/10.1016/j.pep.2024.106521.
Abstract

Plants are often seen as a potent tool in the recombinant protein production industry. However, unlike bacterial expression, it is not a popular method due to the low yield and difficulty of protein extraction and purification. Therefore, developing a new high efficient and easy to purify platform is crucial. One of the best approaches to make extraction easier is to utilize the Extensin Signal peptide (EXT) to translocate the recombinant protein to the outside of the cell, along with incorporating an Elastin-like polypeptide tag (ELP) to enhance purification and accumulation rates. In this research, we transiently expressed Shigella dysenteriae's IpaDSTxB fused to both NtEXT and ELP in both Nicotiana tabacum and Medicago sativa. Our results demonstrated that N. tabacum, with an average yield of 6.39 ng/µg TSP, outperforms M. sativa, which had an average yield of 3.58 ng/µg TSP. On the other hand, analyzing NtEXT signal peptide indicated that merging EXT to the constructs facilitates translocation of IpaDSTxB to the apoplast by 78.4% and 65.9% in N. tabacum and M. sativa, respectively. Conversely, the mean level for constructs without EXT was below 25% for both plants. Furthermore, our investigation into the orientation of ELP showed that merging it to the C-terminal of IpaDSTxB leads to a higher accumulation rate in both N. tabacum and M. sativa by 1.39 and 1.28 times, respectively. It also facilitates purification rate by over 70% in comparison to 20% of the 6His tag. The results of this research show a highly efficient and easy to purify platform for the expression of heterologous proteins in plant.

Introduction

Shigellosis remains a major concern for third-world countries, particularly in Africa and Southeast Asia, with an annual 160 million cases of disease and 600,000 deaths (Bennish and Ahmed 2020). The symptoms are bloody diarrhea, abdominal pain, stomach cramps, and fever which are caused by a gram-negative, rod-shaped bacteria called Shigella (MacLennan et al. 2022). Shigella dysenteriae, among other species such as S. flexneri, S. boydii, and S. sonnei, are the main sources of the disease (Anderson et al. 2016). At the frontline of S. dysenteriae invasion to epithelial cells of human intestines, there is a type III secretion system (T3SS), consisting of four IpaD protein. IpaD is a 37 KDa protein localized at the tip of T3SS, which directly binds to epithelial cells and plays the most pivotal role in invasion (Muthuramalingam et al. 2021). On the other hand, Shigella secretes the Shiga toxin protein, consisting of STxA and STxB domains. While STxB binds to the Gb3 receptor on the epithelial cells, STxA interacts and cleaves the 28s RNA of the 60S subunit of the ribosome, disrupting protein synthesis (Robert and Wiels 2021). Previous research by scientists (Honari et al. 2014; Amani et al. 2023) showed the effectiveness of bacterial expression of the IpaDSTxB chimeric gene in triggering an immune response in mice. Additionally, plant-based IpaDSTxB protein produced in N. tabacum has shown effectiveness against shigellosis in mice, although there remains a challenge of low yield of recombinant proteins (RP) (unpublished data).

Expression and production of RP in plants as green factories such as IpaDSTxB is one of the promising methods that scientists utilize to produce pharmaceutical and therapeutic proteins (Yamamoto et al. 2018). However, unlike bacterial culture, this approach faces several drawbacks, including being time-consuming and costly (Egelrout et al. 2012), low levels of produced RP (Buyel, 2018), technical issues (Watts et al. 2021), a tedious process for obtaining transgenic plants (Shanmugaraj et al. 2020), and challenges with extraction and purification of RP (Matoba et al. 2011). Among these factors, one of the major challenges is the extraction process (Liu et al. 2020). Complete extraction of RP from the plant samples requires the destruction of the whole sample, which is a huge disadvantage because, unlike bacterial culture, regrowing and reaching the
maximum level of the sample for purification is time-consuming (Pam Ismail et al. 2020; Kumar et al. 2021). To address this issue, a method of extracting RP without destroying the sample is transferring and targeting RP to the apoplast and then purifying it using a centrifuge from the media. Signal peptides are the best natural candidates for this purpose (Ishihama et al. 2022). Signal peptides, also known as leader sequences, are short peptides (15–30 amino acids) located at the N-terminal of proteins, which define the destination of the newly secreted protein in the secretion system (Owji et al. 2018). Extensin (EXT) is the most abundant plant cell wall protein, belonging to the class of hydroxyproline-rich glycoproteins (HRGPs) (Zhao et al. 2018). EXT's ability for cross-linking plays an important physiological role in improving the integrity of the plant cell wall, development of pollen growth, and stress and pathogen responses (Moussu and Ingram 2023). The EXT consists of 25 pentapeptide units of S-P-P-P-P, which occur throughout the whole protein. The EXT backbone is synthesized in the ER-bound polysomes, and after hydroxylation of Pro residues in the lumen of the ER, it is transferred to the Golgi apparatus for final glycosylation. Due to its N-terminal 33 amino acid signal peptide, it is directly transferred to the apoplast (Mishler et al. 2021). It seems that utilizing signal peptides, such as EXT, for the extraction of RPs from plants without destroying the sample shows a promising approach to facing challenges in the production of pharmaceutical and therapeutic proteins (Regente et al. 2012).

On the other hand, protein purification using conventional methods such as chromatography columns is difficult, laborious, costly, and often inefficient (Ramos et al. 2019). As a result, the need for an easy-to-use and cost effective method of purification on an industrial scale is a prerequisite (Hedhammar et al. 2006). Elastin-like polypeptide (ELP) is a synthetic biopolymer composed of 10 to 250 repeats of the pentapeptide V-P-G-X-G, with numerous applications in drug delivery (Roberts et al. 2015) and cancer therapy (Varanko et al. 2020). Moreover, ELP is a highly efficient purification tag with over 90% efficiency, as demonstrated by multiple studies (Rodriguez et al. 2016; Phan et al. 2014). ELP purification tag works by exposing hydrophobic residues to the surface and creating an insoluble particle with an increment in temperature (Despanie et al. 2016). Besides its purification capabilities, a synthetic ELP range from 35 to 100 residues in size has been successfully expressed in plant nucleus (Gutiérrez et al. 2013), chloroplasts (Floss et al. 2010), and have been reported to increase RP accumulation (Yang et al. 2021). Although ELP fusions have been shown to increase the accumulation of RPs in plants and aid in purification steps, very little work has been done to investigate the effect of ELP orientation on both purification and accumulation.

The purpose of this study is to develop an improved higher efficient platform for the overexpression of the IpaDSTxB chimeric gene. As a result, we will study the efficiency of the EXT signal peptide in transporting IpaDSTxB to the apoplast of *Nicotiana tabacum* and *Medicago sativa*. Additionally, we will investigate how the orientation of the ELP tag impacts purification rate and RP accumulation.

**Material and Methods**

**Plant and bacterial Materials**

The seeds of both *Nicotiana tabacum* and *Medicago sativa* were obtained from the Seed and Plant Improvement Institute, Iran, and were grown in a greenhouse at 25°C under 16 hours of light and 8 hours of darkness. On the other hand, the *Agrobacterium tumefaciens* strain C58 and *Escherichia coli* strain DH5α were cultured at 28 and 37°C, respectively, in Luria-Bertani (LB) medium containing 100mg/L of Kanamycin and Rifampicin antibiotics (Sambrook and Russell, 2006).
Extensin signal peptide

*N. tabacum* Extensin gene (L38908.1) was used to predict the sequence of the signal peptide by SignalP.5 software (DTU Health Tech). The 63 bp Extensin signal peptide in length was amplified by *Pfu* DNA polymerase with Forward 5' TCCCCGGAGTAGGCCTCC 3' and reverse 5' CCAGGTGTTGCGGTGC 3' primer pairs at 57°C. The amplified sequence was purified from a 1.2% agarose gel, as described by (Zhuo-hua et al. 2000) method and the final product was sent for sequencing.

SOEing PCR and Plasmid Construction

Two ready-to-use pBI-IpaDSTxBELP and pBI-IpaDSTxB vectors were provided by Dr. Alizadeh (unpublished data) for this research. For building the other three constructs, the 563bp IpaDSTxB, 749bp ELP sequences, and 63bp NtEXT sequence were amplified using primers in Table 1–3. In the first step, to build pBI-EXTIpaDSTxBELP, the EXT sequence was merged to the N-terminal of IpaDSTxBELP by *pfu* DNA polymerase at 64°C annealing temperature. The A-tailing process was then carried out after purification. For the second construct, pBI-ELPIpaDSTxB, the IpaDSTxB sequence was amplified and merged to the C-terminal of 749bp ELP at 73°C annealing temperature. Furthermore, for a better comparison, the third construct was built by fusing NtEXT to the N-terminal of ELPIpaDSTxB. Additionally, several restriction sites were introduced into the constructs to increase their versatility. Finally, The amplified EXTELPIpadSTxB, ELPIpaDSTxB, and EXTIpaDSTxBELP sequences were separately cloned into the pUCM T-vector plasmid (BioBasic Inc, Canada), and after verification by sequencing, integrated into the pBI121 via *XbaI* and *SacI* restriction sites, then transformed into DH5α strain according to (Nelson and Fitch 2011).

Table 1
SOEing PCR primer pairs to build EXTELPIpaDSTxB construct, in order to achieve this goal, firstly EXT, ELP, and IpaDSTxB genes were amplified separately by *pfu* enzyme, and after purification 100ng of amplified sequences were mixed with F-ELP and R-IpaDSTxB primers into a reaction to form ELPIpaDSTxB. Finally, by using F-EXT and R-IpaDSTxB primers EXTELPIpaDSTxB were amplified and cloned into pUCM T-vector.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Sequence</th>
<th>Anneling temperature</th>
<th>Restriction sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-EXT</td>
<td>TCTAGAGGATCCTAAAACATGGTCCCCGGAGTAGGCCTCC</td>
<td>74°</td>
<td>Xba I</td>
</tr>
<tr>
<td>R-EXT</td>
<td>GGACCTTATCATCATCATCGCTCGAGGCTGGTTTCCGAAG</td>
<td>75°</td>
<td>Xho I</td>
</tr>
<tr>
<td>F-ELP</td>
<td>CTCGAGGCCTGATGATGATGATAAGGTCC</td>
<td>66°</td>
<td>Xho I</td>
</tr>
<tr>
<td>R-ELP</td>
<td>GAATTCATATCTCCTGGTCAGATATTACC</td>
<td>64°</td>
<td>EcoRI</td>
</tr>
<tr>
<td>F-IpaDSTxB</td>
<td>GTCGACGAGGAGATGATGAATTCCGTACCCAACCAGACG</td>
<td>74°</td>
<td>EcoRI</td>
</tr>
<tr>
<td>R-IpaDSTxB</td>
<td>CATGTTATGATGATCCATGGGAGCTCATTTAAGGAACCGGCTC</td>
<td>73°</td>
<td>Sac I</td>
</tr>
</tbody>
</table>
Table 2
SOEing PCR primer pairs to construct ELPIpaDSTxB. To build this construct ELP and IpaDSTxB sequences were amplified separately by pfu enzyme, and after purification 100ng of amplified sequences were mixed with F-ELP and R-IpaDSTxB primers in a reaction to merge ELPIpaDSTxB. Finally, the amplified sequence were cloned into pUCM T-vector.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Sequence</th>
<th>Anneling temperture</th>
<th>Restriction sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-ELP</td>
<td>CACGCTGTAGCTATCTAGAGCTGATGATGATAAGGTCC</td>
<td>72.5°</td>
<td>Xba I</td>
</tr>
<tr>
<td>R-ELP</td>
<td>GAATTCATATCTCTTCGTCAGACATTACC</td>
<td>64°</td>
<td>EcoRI</td>
</tr>
<tr>
<td>F-IpaDSTxB</td>
<td>GTCGACGAAGGAGATATGAAATCCGTACCACAACGAGGC</td>
<td>74°</td>
<td>EcoRI</td>
</tr>
<tr>
<td>R-IpaDSTxB</td>
<td>CATGTATGATATGACTCCATGGGAGCTCATTAACGGAACCCTC</td>
<td>73°</td>
<td>Sac I</td>
</tr>
</tbody>
</table>

Table 3
SOEing PCR primer pairs to construct EXTIpaDSTxBELP. To build this construct IpaDSTxBELP (Provided by Dr.Alizadeh) and EXT sequence were amplified separately by pfu enzyme, and after purification 100ng of amplified sequences were mixed with F-EXT and R-ELP primers in a reaction to form EXTELPIpaDSTxB. Finally, the amplified sequence were cloned into pUCM T-vector.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Sequence</th>
<th>Anneling temperture</th>
<th>Restriction sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-EXT</td>
<td>TCTAGAGATTCTAATACCTACCCGGAGTGGGCCTTCC</td>
<td>74°</td>
<td>Xba I</td>
</tr>
<tr>
<td>R-EXT</td>
<td>GGACCTTATCATATCATACGCTCCGGTAGCTGGTTTCGGAAG</td>
<td>75°</td>
<td>Xho I</td>
</tr>
<tr>
<td>F-IpaDSTxB</td>
<td>CTTCAGGCTGTAGATGATGATAAGGTCC</td>
<td>66°</td>
<td>Xho I</td>
</tr>
<tr>
<td>R-ELP</td>
<td>CTGCAGAGCGCAGATTCTATCTCTCTCGAGCTCATTAATTACC</td>
<td>73°</td>
<td>SacI</td>
</tr>
</tbody>
</table>

Transgenic Expression

Five vectors, namely: pBI-IpaDSTxB, pBI-IpaDSTxBELP, pBI-EXTIpaDSTxBELP, pBI-ELPIpaDSTxB, and pBI-EXTELPIpaDSTxB, were transformed into Agrobacterium tumefaciens C58 strain and transiently expressed in both N. tabacum and M. sativa according to (Negrouk et al. 2005). After verification by sequencing, resultant strains were incubated overnight at 28°C in 20 mL LB media supplemented with 100 mg/L of kanamycin and rifampicin antibiotics. The overnight incubated recombinant strains were diluted and collected by centrifugation at 5000g for 10 min, and resuspended in 10 mM 2-[N-morpholino]ethanesulfonic acid (MES)-KOH (pH 5.7) containing 10 mM MgCl2, 150 mM acetosyringone, and 50 g/L sucrose (Merck, Darmstadt, Germany). The optical density was adjusted to 0.5 OD. Leaves of N. tabacum and M. sativa were cut and submerged in the Agrobacterium suspension and held for 10 min under 2.5 kPa vacuum conditions. After infiltration, the leaves were washed with sterilized distilled water and kept in darkness for 72 hours at 25°C.

RNA Extraction and qRT-PCR

Total RNA of agroinfiltrated leaves was extracted by the CTAB lithium chloride method according to (Vennapusa et al. 2020). In brief, 100 mg of leaf was homogenized with liquid nitrogen and 1 mL of CTAB buffer containing 100 mM Tris-HCL, 1.4 M NaCl, 20 mM EDTA, and 2% CTAB (Merck, Darmstadt, Germany) was added and incubated for 30 min at 65°C. After centrifugation at 13,000g for 15 min, 600 µL of chloroform was added and
incubated at room temperature. After centrifugation, the supernatant was collected and mixed with 10 M lithium chloride buffer and incubated for 12 hours at 4°C. The resultant pellets were suspended in 0.5% SDS and 500 µL of chloroform and incubated on ice for 10 min. After centrifugation, the pellet was resuspended and washed with cold isopropanol and 75% ethanol, respectively. Contaminated genomic DNA in the total RNA, was degraded with DNase (Thermo Fisher Scientific) and both RNA concentration and purity were determined using a NanoDrop® ND-1000 spectrophotometer (ThermoFisher Scientific, DE), And 1.2% agarose gel electrophoresis. To generate cDNA, one microgram of total extracted RNA was used, and cDNA were generated by cDNA synthesis kit (BioBasic Inc, Canada). The qRT-PCR was performed in quadruplicate with IpaDSTxB forward 5’ CGGTGTCTTTACACCCGT 3’ and reverse 5’ CAGGTTCCAGCGGTAGTA 3’ primer pairs in 40 cycles with an annealing temperature of 60°C using QIAGEN’s real-time PCR system (Rotor-Gene Q). The β-actin gene was used as an internal control, and the data analysis of gene expression was carried out using REST software.

**Protein Extraction and ELISA**

Apoplastic proteins were extracted according to (Jaswanthi et al. 2019). In brief, agroinfiltrated leaves were carefully rolled into clean parafilm tape and fixed in the syringe barrel and placed into 50 mL centrifuge tubes, and centrifuged at 7000g for 6 min at 4°C. The remaining leaves were used for symplastic protein extraction. Total and symplastic proteins were extracted by adding 1 mL of protein extraction buffer (25 mM sodium phosphate, 1 mM EDTA, 0.1% Triton X-100, 100 mg/mL sodium ascorbate, and 0.3 mg/mL phenylmethylsulfonyl fluoride [PMSF]) to agroinfiltrated homogenized leaves and centrifuged at 14000g for 10 min at 4°C. Protein concentration was determined by Bradford assay using Bovine serum albumin. Finally, concentration of TSP, apoplastic, and symplastic recombinant proteins were determined using indirect ELISA assay according to (Hajibehazd et al. 2016) method.

**Protein purification**

Purification of pBI-IpaDSTxB protein was carried out by utilizing a HisTrap FF 1mL column (GE Healthcare, Sweden) and Amicon Ultra-4 10K centrifugal filter device (Merck Millipore). For the purification of other four constructs, the method described by (Sarvestani et al. 2021) was used in quadruplicate.

**Statistical analysis**

The experiment was designed based on a completely randomized factorial experiment, and all the data analyzed by SAS 9.4 software. Mean value compared by latest significance difference (LSD test) and P-value < 0.05 considered significant.

**Results**

**Plasmid construction**

Results of T-vector digestion of pUCM-EXTELPIpaDSTxB, pUCM-ELPIpaDSTxBEXT, and pUCM-ELPIpaDSTxB by *Xbal* and *SacI* restriction enzymes showed a distinct band on 1.2% agarose gel at around 1.5kbp (Fig. 1a). For Colony-PCR, 50ng of each vectors and IpaDSTxB primer pairs were used in PCR reactions to amplify a 550bp sequence of IpaDSTxB (Fig. 1b). Furthermore, digestion of pBI constructed vectors with *Xbal* and *SacI* restriction enzymes showed two clear bands, one backbone of pBI121 over 3Kbp marker and the the other
representing the coloned sequence around 1.5kb (Fig. 1c). Finally, Sequencing analysis of the constructed plasmids has also shown the accuracy of the amplified fragments (Fig. 2).

**qRT-PCR analysis**

To measure recombinant transcripts levels, qRT-PCR was carried out on agroinfiltrated leaves of *N. tabacum* and *M. sativa*. β-Actin was used as the reference gene and agroinfiltrated leaves of non-recombinant C58 were used as the negative control. The results confirmed the expression of all constructs in both plants after 72 hours. The most highly expressed constructs were pBI-EXTIpDSTxBELP and pBI-IpaDSTxBELP in *N. tabacum*, with mean levels of 7.92 and 6.63, respectively. In contrast, the lowest mRNA concentration of was for pBI-IpaDSTxB with an expression level of 2.57 in *M. sativa*. On the other hand, the highest expression level was for pBI-ELPIpDSTxB and pBI-EXTIpDSTxBELP at 4.86 and 4.72, respectively. Finally, the average expression level of all constructs in *N. tabacum* was 6.32, and 3.89 in *M. sativa*, showing that the expression system of *N. tabacum* was 1.62 times stronger than *M. sativa* (Fig. 3).

**ELP orientation and Purification**

To investigate the effect of ELP orientation on RP accumulation, the ELP tag has been merged into both the N-terminal and C-terminal of IpDSTxB. The results indicate that fusing ELP in both the N-terminal and C-terminal, compared to pBI-IpaDSTxB, can boost RP production by 1.5 and 2.23 times in *N. tabacum* and *M. sativa*, respectively. Specifically, the highest accumulation levels were found in *N. tabacum* leaves for both EXTIpDSTxBELP and IpaDSTxBELP, with a mean accumulation of 7.99 ng/µg compared to 5.74 ng/µg of the EXTELPIpDSTxB and pBI-ELPIpDSTxB constructs. In *M. sativa*, the average accumulation level for both EXTIpDSTxBELP and IpaDSTxBELP was 4.52 ng/µg, while this number for EXTELPIpDSTxB and ELPIpDSTxB was 3.51 ng/µg. These results show that merging ELP to the C-terminal of IpaDSTxB leads to higher accumulation compared to N-terminal integration in both plants, boosting levels by 1.39 and 1.28 times in *N. tabacum* and *M. sativa*, respectively. This means that fusing ELP to the C-terminal of recombinant genes has a significant impact on overexpression and accumulation rate (Fig. 4).

Furthermore, we compared the purification rate by studying the efficiency of ELP tag against 6His tag. The results suggest that using the ELP tag facilitates the purification of all chimeric ELP constructs by over 70% from 1 µg of total soluble protein in both plants. The results also indicate that the different orientation of the ELP tag has no significant effect on the purification rate in all samples. The mean purification level for all four constructs tagged with ELP in *N. tabacum* was 72.92% (5 ng/µg from 6.87 ng/µg TSP) and for *M. sativa* was 72.34% (2.91 ng/µg from 4.02 ng/µg TSP) (Fig. 5). In contrast, the purification rate of utilizing the 6His tag in *N. tabacum* and *M. sativa* was 21.73% (0.97 ng/µg from 4.49 ng/µg TSP) and 20.3% (0.36 ng/µg from 1.82 ng/µg TSP), respectively. Thus, a 3.45 times higher purification rate was observed by using the ELP purification tag compared to the 6His tag.

**Total soluble protein, Apoplastic, and Symplastic protein**

Total soluble protein of agroinfiltrated leaves of *N. tabacum* and *M. sativa* was analyzed by indirect ELISA assay in quadruplicate. Based on the results, In *N. tabacum* leaves, the average level of recombinant protein accumulation was measured at 6.39 ng/µg TSP, while in *M. sativa* it was found to be 3.58 ng/µg TSP (Fig. 4). This means that around 0.63% and 0.35% of TSP of *N. tabacum* and *M. sativa* are recombinant protein, which shows a 1.78-fold difference in folding between these two plants. Among the vectors, the highest RP
concentration was observed in *N. tabacum* leaves with an average of 8.17 ng/µg for pBI-EXTIpaDSTxBELP, and the lowest concentration was for pBI-IpaDSTxB with an average of 1.82 ng/µg in *M. sativa*.

Results of Apoplastic protein analysis showed that the EXT signal peptide has a significant role in the accumulation of RP in the apoplast of both plants. The highest apoplastic RP accumulated in agroinfiltrated leaves of *N. tabacum* was detected with pBI-EXTELPIpaDSTxB and pBI-EXTIpaDSTxBELP with an average of 82.23% (5.04 ng/µg from 6.21 ng/µg TSP) and 73.85% (6.03 ng/µg from 8.17 ng/µg TSP), respectively (Fig. 6). On the contrary, the lowest concentration was observed with pBI-IpaDSTxB with an average of 25.69% (1.24 ng/µg from 4.49 ng/µg TSP). On the other hand, we observed a similar pattern of Apoplastic RP accumulation in *M. sativa*. While around 23.81% (0.55 ng/µg from 1.82 ng/µg TSP) and 23.67% (0.66 ng/µg from 2.83 ng/µg TSP) of RP of agroinfiltrated leaves of *N. tabacum* with pBI-IpaDSTxB and pBI-ELPIpaDSTxB were accumulated in the apoplast, this number for pBI-EXTELPIpaDSTxB and pBI-EXTIpaDSTxBELP was approximately 66.2% (2.84 ng/µg from 4.2 ng/µg TSP) and 65.5% (2.9 ng/µg from 4.58 ng/µg TSP), which was almost three times higher than other constructs.

Furthermore, the average apoplastic RP concentration showed that there was a significant difference in efficiency of EXT signal peptide in both plants. While this number for constructs harboring NtEXIT was 78.4% in *N. tabacum*, this number for *M. sativa* was detected as 65.9%, which means that EXT signal peptide’s efficiency was 1.18 times higher in *N. tabacum*. Additionally, an average of 24.9% and 23.3% of RP accumulated in the apoplast by constructs without NtEXIT in both *N. tabacum* and *M. sativa* showed that merging EXT to the N-terminal of RP boosts translocation of RP to the apoplast by 3.13 and 2.82 times, respectively.

**Discussion**

In the present study, our first goal was to investigate the efficiency of the NtEXIT signal peptide in transferring and locating IpadSTxB to the apoplast of *N. tabacum* and *M. sativa*. Furthermore, we studied the efficiency of ELP purification tag against conventional 6His tag for purification purposes. Finally, we investigated the effect of both N and C-terminal orientations of ELP tag on the accumulation of IpaDSTxB. As a result, we managed to successfully build three gene constructs namely; pBI-ELPIpadSTxB, pBI-EXTELPIpadSTxB, pBI-EXTIpadSTxBELP and after verifying them by sequencing and digestion, we transiently expressed them along with pBI-IpadSTxB and pBI-IpaDSTxBELP vectors in *N. tabacum* and *M. sativa* leaves by utilizing *A. tumefaciens* C58 strain and vacuum-based agroinfiltration method.

The result of qRT-PCR showed that there was a significant difference in the expression level of RP between *N. tabacum* and *M. sativa*, and the expression level of *N. tabacum* was 1.62 times higher than *M. sativa*. This difference could be related to our codon usage optimization of all sequences with *N. Tabacum* genome. Due to the role of codon usage in mRNA stability (Presnyak et al. 2015) and transcription efficiency (Liu et al. 2021), it is possible that *N. tabacum*’s high expression level could be related to the highest mRNA stability and efficiency of its transcription machinery. This result was aligned with the reports of (Bellucci et al. 1997, 2007). In both of these reports, the expression of Zeolin and γ-Zein (optimized with *N. tabacum*) in *N. tabacum* was 1.23 and 1.86 times higher than *M. sativa*, respectively. Additionally, the number of transformed t-DNA (Pereman et al. 2019) into the nucleus and the effect of timing on the degradation of foreign t-DNA (Gelvin 2021) could be another possible explanation for the difference in expression among constructs and plants.
Consistent with the qRT-PCR findings, the mean concentration level of all RPs in TSP shows that \textit{N.tabacum} has a 1.78 advantage in producing RP over \textit{M.sativa}. This significant difference is not only related to codon usage bias but also due to overall gene expression differences (Zhang et al. 2020), the effect of Post-Transcriptional Gene Silencing (Liu and Timko, 2022), and protease level (Pillay et al. 2014). For example, it has been demonstrated that RNA-dependent RNA polymerase 6 (RDRP6) detects and eliminates excessive recombinant transcripts in \textit{A.thaliana} (Beclin et al. 2002). This is related to the production of unpolyadenylated mRNAs due to issues like premature transcription termination, incomplete elongation, or readthrough beyond the usual transcription end site. These abnormal mRNAs may impact gene expression and protein production (Feng et al. 2022). On the other hand, (Ziegelhoffer et al. 1999) has shown that due to higher protease levels in \textit{M.sativa} and \textit{S.tuberosum}, the expression of bacterial cellulase E2 and E3 genes was 0.69 times higher in \textit{N.tabacum}. Aligned with their findings, the study of (Grosse-Holz et al. 2018) proved that a higher expression level of protease inhibitors such as NbPR4 and NbPot1 had a positive and significant impact on the accumulation of RPs in \textit{N.benthamiana}. Consequently, it is not unexpected that due to various mechanisms like RNA silencing and proteolysis, agroinfiltrated leaves from different plant species would yield varying levels of RPs. Moreover, based on the results, the highest concentration of RPs in both plants was detected in agroinfiltrated leaves of pBI-IpaDStxBELP and pBI-EXTIpaDStxBELP. This means that merging ELP to the C-terminal of IpaDSTxB had a positive effect on the accumulation of RPs in comparison to N-terminal merging. Although the exact underlying reason is unknown, this could be related to the reduction of protein misfolding (Johnston and Samant, 2021) and improving protein stability (Brich-Machin et al. 2004). Our analysis suggests that merging ELP to the N-terminal of IpaDSTxB may have a significant impact on increasing RP misfolding, leading to aggregation and protein degradation by proteases. However, the results also show fusing ELP to both sides has a beneficial impact on protein accumulation and concentration in comparison to pBI-IpaDSTxB. Furthermore, studies (Scheller et al. 2006; Wu et al. 2006) proved that fusing ELP to RP has a significant positive effect on protein accumulation by increasing protein’s stability. Additionally, (Patel et al. 2007) observed that merging ELP to interleukin-4 in transiently expressed \textit{N.tabacum} leaves can increase the accumulation of IL4-ELP 19 times higher than IL4-6His. Furthermore, our over 70% ELP purification rate was similar to the 95% purification rate of bacterial-expressed Human Epidermal Growth Factor (hEGF) by (Sarvestani et al. 2021) and 76% of IL-4 by (Patel et al. 2007). This finding showed and proved that not only is ELP one of the easiest, fastest, and most efficient purification tags for a large amount of RP production but also both its orientations have a positive impact on protein accumulation (Conley et al. 2009).

Signal peptides are considered a powerful tool in targeting and translocating proteins to other organelles and the extracellular space (Castilleux et al. 2021). One of the key proteins in the plant cell wall is a glycoprotein called Extensin, with a 23 to 32 residue signal peptide in the N-terminal (Ding et al. 2020). Our study showed that utilizing EXT can target around two-thirds of RPs to the apoplast in both \textit{N.tabacum} and \textit{M.sativa}. This aligns with previous studies indicating that only 18 residues of \textit{N.plumbaginifolia} signal peptide are necessary in transferring the NPTII protein to the MS media (De Loose et al. 1991). Furthermore, studies by (Goojani et al. 2013) on KDEL and EXT signal peptides showed that using both signal peptides increased the expression of RPs from 5 ng/µg TSP to 6.9 ng/µg TSP. This increase in expression, which we also found in our analysis, could be related to the negative feedback regulation phenomena (Zhao et al. 2020). While it has been known that the accumulation of protein in the cell could result in protein aggregation and reduction of expression, merging EXT to RP and its translocation to the apoplast could maintain stable transcription, leading to a higher expression level of protein (Schramm et al. 2020).
Conclusion

In this research, we carefully analyzed the effect of NtEXT signal peptide and the orientation of ELP purification tag in the expression and accumulation of *S. dysenteriae* IpaDSTxB recombinant protein in both *N. tabacum* and *M. sativa* as an alternative method to not only increase accumulation and purification rates, but also to make the whole process easier on an industrial scale. First of all, our findings suggest that fusing NtEXT signal peptide has a positive and significant impact on transferring IpaDSTxB to the apoplast, and it may also have a possible effect on increasing RP. Secondly, we realized that merging ELP purification tag to the C-terminal of RP can boost protein accumulation levels, as well as increase purification ability in comparison to conventional purification methods. Finally, we introduced an optimized platform of transiently expressed RPs in *N. tabacum* by fusing NtEXT and ELP to the N and C-terminal to our RP, and we believe that by conducting further studies, this platform has the potential for commercialization.

Declarations

Competing interests

The authors declare no competing interests.

Conflict of Interest

The authors declare that they have no conflict of interests.

Author contributions

AMS: Writing-original draft preparation, Writing-review and editing, Visualization, Methodology, Formal analysis and investigation and Validation HA: Project administration, supervision, Conceptualization, Writing-review and editing, Validation and Resources. All authors have read and approved the final manuscript.

Data availability

The data that support the findings of this study are available from the corresponding author upon request.

References


Figures

Figure 1

Analysis of digestion by restriction enzymes and Colony PCR pairs on 1.2% agarose gel, (A) recombinant T-vector Digestion by \textit{XbaI} and \textit{SacI} restriction enzymes, Lane 1,2,3: 1.4Kbp band of EXTEL\textit{IpaDSTxB}, \textit{EXTIpaDSTxBELP}, and \textit{ELPIpaDSTxB}, Lane 4,8,10: DNA Marker (DM2300, SmoBio, Hsinchu, Taiwan) (B) Colony PCR confirmation by amplifying 550bp \textit{IpaDSTxB} gene Lane 5,6,7: 550 bp band of EXTEL\textit{IpaDSTxB}, \textit{EXTIpaDSTxBELP}, and \textit{ELPIpaDSTxB}, Lane 9: Negative control (C) result of recombinant pBI digestion by \textit{XbaI} and \textit{SacI} restriction enzymes Lane 11,12,13: 1.5Kbp band of EXTEL\textit{IpaDSTxB}, \textit{EXTIpaDSTxBELP}, and \textit{ELPIpaDSTxB}. 
Figure 2

Schematic of utilized constructs (A) pBI-IpaDSTxB (B) pBI-ELPpBI-DSTxB (C) pBI-EXTpBI-DSTxBELP (D) pBI-IpaDSTxBELP (E) pBI-EXTpBI-DSTxBELP. Both (A) and (D) constructs were provided by Dr. Alizadeh.
Figure 3

Relative expression levels of IpaDSTxB, EXTELPIpaDSTxB, ELPIpaDSTxB, IpaDSTxBELP, and EXTIpDSTxBELP mRNA in both *N. Tabacum* and *M. Sativa*, 72 hours after agroinfiltration.
**Figure 4**

Total soluble protein of recombinant IpaDSTxB, EXTELPaaDSTxB, ELPLpaDSTxB, IpaDSTxBELP, and EXTIpaaDSTxBELP from N. Tabacum and M. Sativa, 72 hours after agroinfiltration.

![Bar chart showing purification rate of recombinant proteins in N. Tabacum and M. Sativa.](image)

**Figure 5**

Protein purification rate of recombinant IpaDSTxB, EXTELPaaDSTxB, ELPLpaDSTxB, IpaDSTxBELP, and EXTIpaaDSTxBELP using ELP purification tag against conventional 6His tag in both *N. Tabacum* and *M. Sativa*, 72 hours after agroinfiltration from 1µg of TSP.
Figure 6

Apoplastic protein of recombinant IpaDSTxB, EXTELPIpaDSTxB, ELPIpaDSTxB, IpaDSTxBELP, and EXTIpaDSTxBELP from N. Tabacum and M. Sativa, 72 hours after agroinfiltration.