JOURNAL OF APPLIED Plant Biology
2024, Vol2. No1.pp 11-1117
DOI: 10.22034/JAPB.2024.104

Construction of Mambalgin-1 gene cassette for transient expression in apoplastic space

Ghaffar Khezri[1]*, Iman Yousefi Javan[1]

[1]Department of Plant Production, Faculty of Agriculture, University of Torbat Heydarieh, Torbat Heydarieh, Iran

*Correspondence : Ghaffar Khezri, **E. mail:** gh.khezri@torbath.ac.ir, khezrighaffar@gmail.com

Abstract

Mambalgine-1 is a protein that acts by blocking ASICs channels in neurons as a potent analgesic such as morphine while having no adverse effects of morphine, such as addiction and respiratory distress. In this study, in order to produce inexpensive and scalable Mambalgin-1 protein, recombinant PVX-Mambalgin-1 viral vector was designed and optimized for transient expression in apoplastic space of *Nicotiana benthamiana*. To compensate for the lack of methionine at the beginning of Mambalgin-1 and also secretion of this protein into apopalstic space Glucanase inhibitor protein 2 (GIP2) signal peptide was added to the N-terminal of Mambalgin-1. The 6xhis-tag was added to the C-terminal of Mambalgin-1 for downstream processes and protein purification. Probability of signal peptide cleavage and secretory of Mambalgin-1 assessed by SignalP.5 server. The desired protein was reverse translated into nucleotide sequences based on *N. benthamiana* Codon Usage from Kazusa Database; then mRNA destabilizing sequences, polyadenylation signal and hidden stop codons were modified using Visual Gene Developer 1.9 and Mega4 softwares. CAI and GC were kept as high as possible. ClaI/SalI restriction enzyme sites were added at the 3' and 5' end of the construct. The results showed that the GIP2 signal peptide would most likely (0.96) be cleaved from Mambalgin-1 and Mambalgin-1 expression will be highly secretory (0.998) into the apopalstic space after expression. After optimization, cloning and transfection of Agrobacterium, the existence and transformation of recombinant PVX-Mambalgin-1 vector was confirmed by colony PCR

Keywords: Mambalgin-1, Glucanase inhibitor protein 2, In silico, Transient expression, Apoplastic space

1.Introduction

Mambalgin-1 is a 57-amino acid peptide discovered from Mamba snake venom, having a potent analgesic effect like morphine by blocking Acid Sensing Ion Channels (ASICs) in central and peripheral nervous system (1). This peptide has no side effect like morphine such as naloxone susceptibility, respiratory distress and much less tolerance than morphine in mice. (1-3). Most of the researches that have been performed on Mambalgin-1 has focused more on its structure and relationship with ASICs (3-5). Plants have significant advantages compared to hosts producing recombinant proteins, such as scalability, low cost, non-contamination with human pathogens, and the ability to perform post-translational processes such as glycosylation (6-10).

Transient expression systems for recombinant protein production, have advantages over permanent transgenic plants such as simplicity, lower cost, omitting tissue culture and regeneration costs and ease of large-scale plans (11-14). In transient expression systems, the use of hybrid viral vectors based on plant viruses and *Agrobacterium* T-DNA has become very attractive due to the high recombinant protein production in a short time. Until now, various therapeutic proteins for influenza, hepatitis, plague, smallpox, West Nile virus, etc. have been produced using transient expression systems based on plant viral vectors (11, 12, 15, 16). In this study, Mambalgin-1 gene cassette components were designed, codon-optimized and constructed in the hybrid potato virus X (PVX) vector for *Nicotiana benthamiana*, these laid a foundation for the transient expression and purification of recombinant Mambalgin-1 in the future.

2. Materials and methods

Design of gene construct for recombinant protein components

To compensate for methionine at the beginning of Mambalgin-1 protein, Secretory and removable Glucanase inhibitor protein 2 (GIP2) signal peptide derived from (amino acid sequence: MASSCCLHAILLCSLLFITSTTA) (Uniprot accession number: P0DO21.1) of *Nicotiana benthamiana* (17) added to N-terminal of Mambalgin-1. The probability of cleavage of the GIP2 signal peptide from the Mamblagin-1 protein was determined by the SignalP-5.0 server (18) (http://www.cbs.dtu.dk/services/SignalP/). Also, Hexa histidine (6xHis-tag) tag for nickel or cobalt affinity purification added to C-terminal of Mambalgin-1.

Codon optimization of desired protein and construction of the vector

The desired protein obtained in previous step was reverse translated from amino acid to DNA. Codon optimization was performed with Visual Gene Developer Software Ver. 1.9 (19). In foregoing process of codon optimization the following factors were considered such as: i) try to keep up the Codon Adaptation Index (CAI) of nuclear-encoded genes of *Nicotiana benthamiana* (20, 21) ii) try to keep up of GC content of optimized sequence iii) Removing mRNA destabilizing sequences iv) detection of restriction enzyme sites and removing them v) Modification of polyadenylation signals vi) addition of stop codon at the 3′ end and modification of the hidden stop codon between sequence codons viii) addition of *Cla*I and *Xba*I restriction sites at the 5′ and addition *Sac*I and *Sal*I,restriction sites at the 3′ of the gene cassette. The optimized sequence for expression in *Nicotiana bentamiana* was chemically synthesized and cloned in *Cla*I/*Sal*I restriction sites of PVX-GW vector under the control of duplicated PVX-coat protein subgenomic promoter (CPP).

Transformation of Agrobacterium tumefaciens by plant expression vector

Transformation of PVX-Mambalgin-1 into *Agrobacterium tumefaciens* strain GV3101 was done by standard freeze-thaw protocol (Höfgen and Willmitzer, 1988). Transformed *Agrobacterium* were selected on selective medium containing 100 µg/mL rifampicin (RIF) and 50 µg/mL kanamycin (KAN), incubated for 72 hours at 28°C. Transformed Agrobacterium was confirmed by colony PCR using forward and reverse primers (Table 1, were designed with the Oligo Primer Analysis Software version 7 (22).

Table 1: Sequences of primers used in the colony PCR

Primer Name	Oligo sequence 5´ → 3´	Nomber of Bases	Fragment length (bp)
Forward	CACCAGCTAGCATCGATTC (PVX-GW, ClaI, XbaI)	19	300
Reverse	GGTCGACGAGCTCTTATTAATG (PVX-GW, Sal, Sac, Stop Codons, Histag)	22	

3. RESULTS

Design of recombinant protein components

Figure 1 shows the recombinant protein components, including signal peptide, Mambalgin-1 and His-tag sequences and their arrangement. As can be seen in figure 2 (output of results from SignalP-5.0 server) the prediction of the precise cleavage between the signal peptide and Mambalgin-1 is highly probable (0.9652) and after the signal peptide is cleaved by the signal peptidase I enzyme, Mambalgin-1 will be secreted into the secretory space (Probability: 0.9984).

Figure 1. The amino acid sequence of the desired protein and its components comprising the Glucanase inhibitor protein 2 (GIP2) peptide signal at the N-terminal and the His-tag sequence at the C-terminal of Mambalgin-1. UniProt accession numbers of signal peptide and Mambalgin-1 are P0DO21.1 and P0DKR6 respectively.

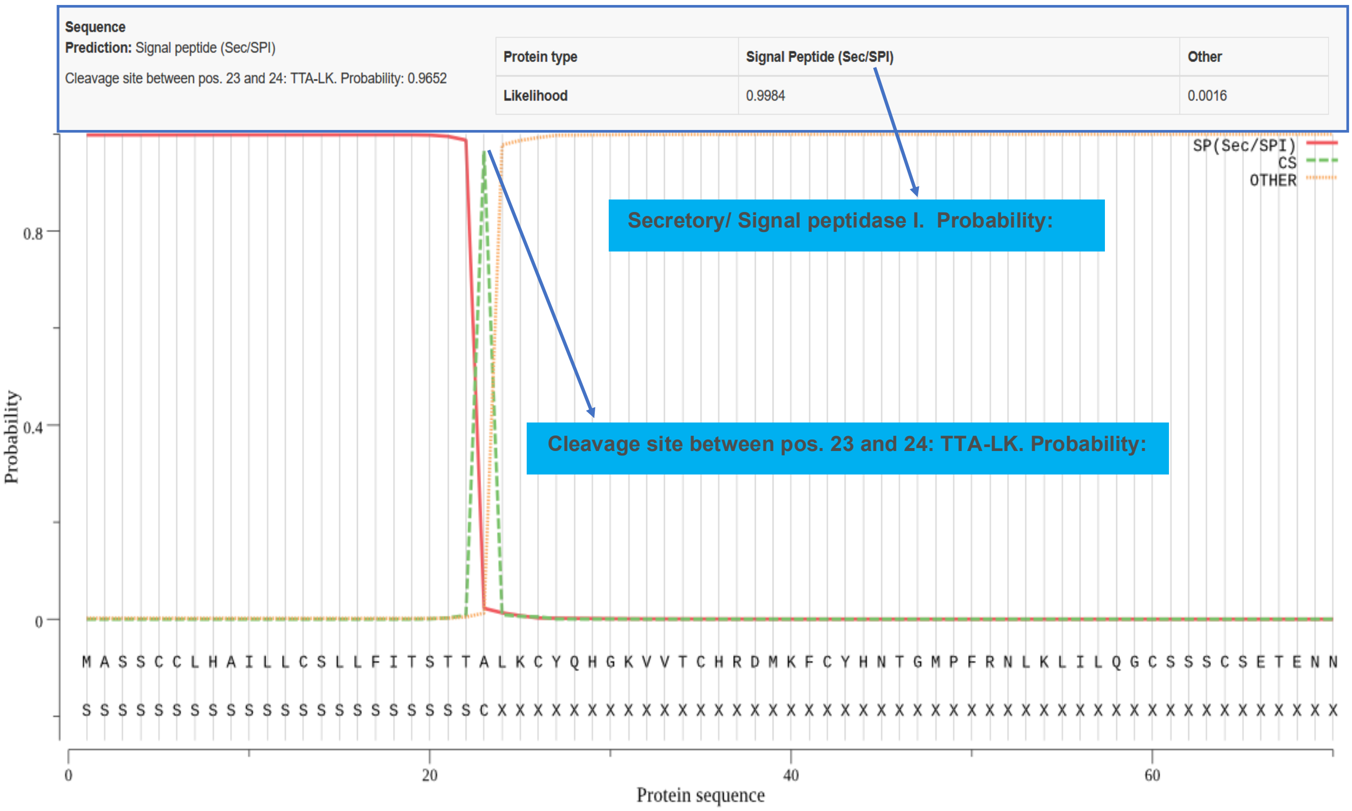

Figure 2. Output of prediction of cleavage site between the signal peptide (Glucanase inhibitor protein 2 (*Nicotiana benthamiana*)) and Mambalgin-1 in desired protein sequence from SignalP-5.0 Server; Cleavage site between position 23 and 24: TTA-LK. Probability: 0.9652. The signal peptide is secretory and will be cleaved by the signal peptidase I. Probability: 0.9984

Codon optimization of desired construction of the protein and vector

Figure 3 shows the nucleotide modification in the nucleotide sequence resulting from the reverse translated amino acid sequence for efficient and effective expression in *Nicotiana bentamiana*. Sequences "ATAGAT" (nucleotides 239-244), "TTTTT" (nucleotides 59-63) and "GGTAAG" (nucleotides 100-105) were modified as mRNA destabilizing sequences (23, 24). The "AATAAT" (nucleotides 59-63) sequence was identified and modified as a polyadenylation signal (25). The hidden stop codons found between the main codons were identified and changed for "TAA" (nucleotides 84, 102, 144, 171, 219, 222, 246 and 249), "TGA" (nucleotides 208, 213 and 237) and "TAG" (121, 162 and 240). The other codon changes seen in figure 3 were due to the higher GC content and codon optimization. Restriction enzyme sites were added to the 5' and 3' ends. As can be seen in figure 4, after optimization, the GC content and Nc increased from 31% and 12.5 to 41% and 19.5, respectively. The codon adaptation index (CAI) also declined from 1 to 0.8.

Amino acid			ClaI			XbaI		M	A	S	S	C	C	L	H	A	I	L				
Back translated								ATG	GCT	TCT	TCT	TGT	TGT	CTT	CAT	GCT	ATT	CTT				
Optimal			ATCGAT			TCTAGA		ATG	GCA	TCC	TCC	TGC	TGT	TTA	CAT	GCT	ATT	CTG				
Amino acid	L	C	S	L	L	F	I	T	S	T	T	A	L	K	C							
Back translated	CTT	TGT	TCT	CTT	CTT	TTT	ATT	ACT	TCT	ACT	ACT	GCT	CTT	AAG	TGT							
Optimal	TTG	TGC	AGC	CTG	TTG	TTC	ATT	ACT	TCT	ACC	ACA	GCA	CTC	AAG	TGT							
Amino acid	Y	Q	H	G	K	V	V	T	C	H	R	D	M	K	F							
Back translated	TAT	CAA	CAT	GGT	AAG	GTT	GTT	ACT	TGT	CAT	AGA	GAT	ATG	AAG	TTT							
Optimal	TAT	CAA	CAT	GGG	AAG	GTT	GTT	ACT	TGT	CAT	CGG	GAT	ATG	AAG	TTC							
Amino acid	C	Y	H	N	T	G	M	P	F	R	N	L	K	L	I							
Back translated	TGT	TAT	CAT	AAT	ACT	GGT	ATG	CCT	TTT	AGA	AAT	CTT	AAG	CTT	ATT							
Optimal	TGT	TAT	CAC	AAC	ACT	GGT	ATG	CCG	TTC	AGA	AAT	CTC	AAA	CTC	ATT							
Amino acid	L	Q	G	C	S	S	S	C	S	E	T	E	N	N	K							
Back translated	CTT	CAA	GGT	TGT	TCT	TCT	TCT	TGT	TCT	GAA	ACT	GAA	AAT	AAT	AAG							
Optimal	CTT	CAA	GGT	TGT	TCT	TCT	TCT	TGT	TCA	GAG	ACA	GAA	AAC	AAC	AAG							
Amino acid	C	C	S	T	D	R	C	N	K	H	H	H	H	H	H							
Back translated	TGT	TGT	TCT	ACT	GAT	AGA	TGT	AAT	AAG	CAT	CAT	CAT	CAT	CAT	CAT							
Optimal	TGT	TGT	TCT	ACC	GAC	AGG	TGC	AAC	AAG	CAT	CAT	CAT	CAT	CAT	CAT	GAGCTC		GTCGAC				
																SacI		SalI				

Figure 3. Nucleotide sequence comparison between back translated of desired protein and *Nicotiana benthamiana* plant optimized Mambalgin-1.

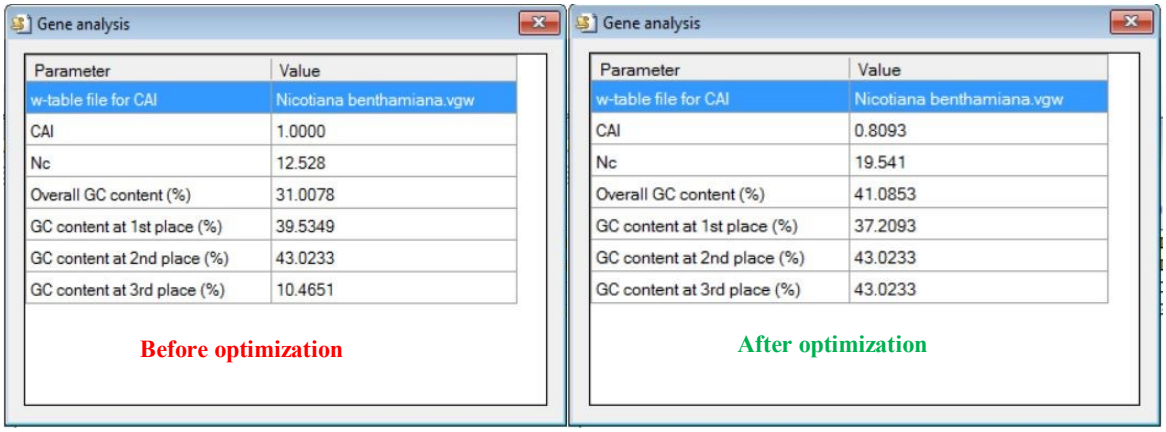

Figure 4. Status of codon optimization parameters before and after optimization obtained from Visual Gene Developer 1.7 software, CAI: codon adaptation index, Nc: number of effective codons.

Construction of recombinant vector and confirmation of its presence in Agrobacterium

Figure 5 shows the gene cassette entry under the CP promoter and upstream of the NOS-terminator transcription terminator in the PVX-GW vector. According to the results of colony PCR using specific primers (table 1) in figure 6, the presence of recombinant vector containing gene cassette (signal peptide +mature Mambalgin-1 + 6 His tag) was confirmed.

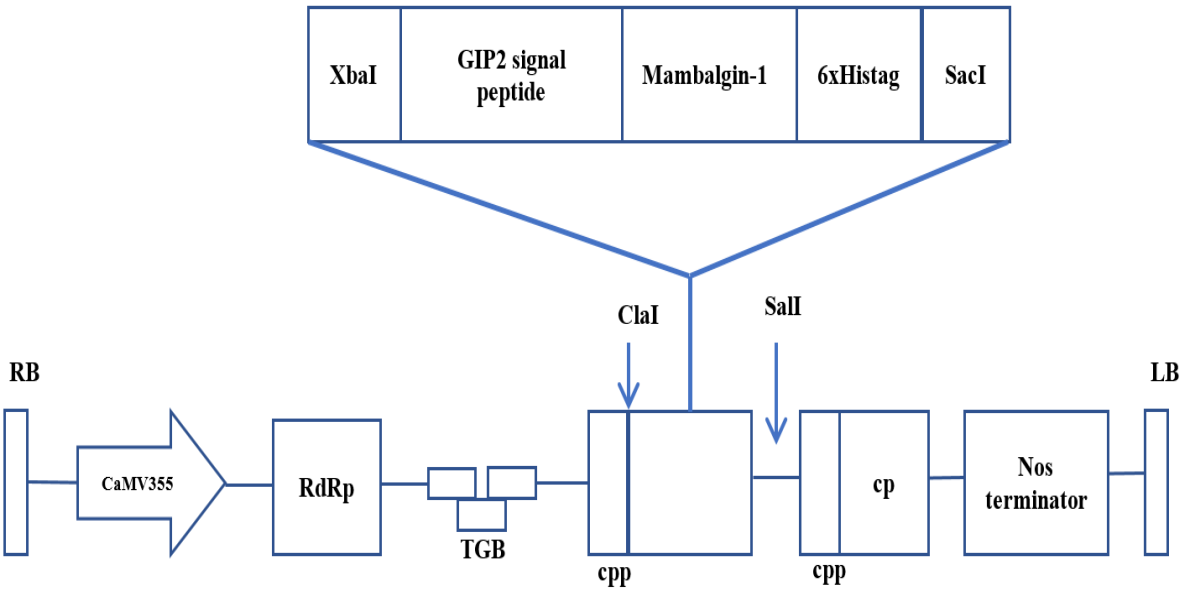

Figure 5. Schematic diagram of the constructed PVX–Mambalgin-1 vector: The synthetic and optimized gene was inserted into the ClaI/SalI sites of PVX-GW vector under the control of duplicated PVX-coat protein subgenomic-promoter (CPP). RdRp: RNA dependent RNA polymerase, TGB: Triple-gene-block.

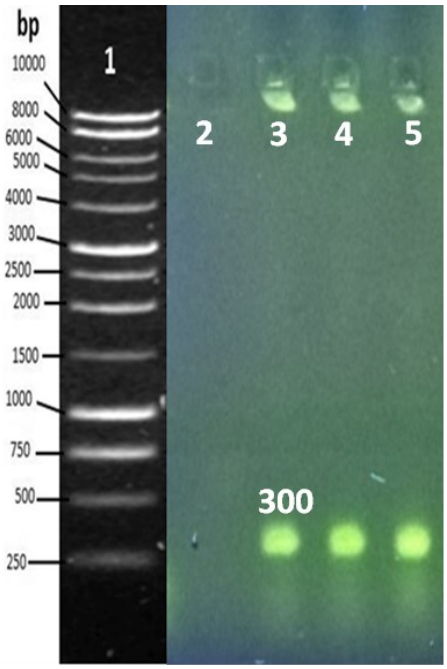

Figure 6. Colony PCR analysis; lane 1: 10kb DNA ladder, lane 2: non-transformed *Agrobacterium tumefaciens* GV3101, lanes 3, 4 and 5 positive clones

4. Discussion

The presence of methionine amino acids is essential for translation at the N-terminal of protein and, as can be seen in figure 1, the Mambalgin-1 protein lacks methionine and the GIP2 signal peptide was used to compensate for it. On the other hand, after expression, the possibility of separation and cleavage of the signal peptide should be investigated to produce the signal-free mature recombinant protein. The SignalP-5.0 server was used to investigate the probability of cleavage and detachment. In the results obtained in figure 2, the probability of GIP2 signal peptide cleavage is very high (0.9652) and can release mature Mambalgin-1. Given the nature of the GIP signal to be secreted, the results also indicate that the highly probable. Mambalgin-1 protein will become a secretory protein and will move into the apoplastic space after expression. The secretion of recombinant protein into the extracellular space is an advantage that does not require complex purification (26). Polyadenylation signals can cause undesirable processing of mRNA (27, 28), and according to this logic, the "AATAAT" (nucleotides 59-63) sequence was identified and modified as a polyadenylation signal (25). Given that there is a direct relationship between the GC content and the expression level of the recombinant protein, after optimization, the GC content increased from 30% to 41% (figure 4). One of the most important factors for effective and efficient expression of recombinant protein in transient expression systems is the codon adaptation index (CAI) in the host plant (29, 30). As can be seen in figure 1, the CAI decreased from 1 to 0.8 due to the balancing of all optimization parameters. In addition to balancing all the optimizations and in the optimization process, it has attempted to have a CAI above 0.8 and this value shown in figure 1 is scientifically acceptable (31). Transfer and presence of recombinant vector in *Agrobacterium tumefaciens* GV3101 was confirmed by colony PCR assay and this Agrobacterium was stored and prepared for agroinfiltration of *Nicotiana benthamiana* leaves to high levels of transient expression of Mambalgin-1 protein.

5. Conclusion

In silico studies are required before any effort to produce a recombinant protein. Our study revealed that addition of GIP2 signal peptide to mature mambalgin-1 protein guaranties its efficient production and secretion into the apopalstic space while avoiding Methionine addition to its N-terminal may affect structure and function of the mature protein. Codon optimization of reverse translated gene construct and removing polyadenylation signals and hidden stop codons results in maximum expression of heterologous protein in the host is aimed to produce mambalgin-1 protein as an efficient biofactory.

Conflict of Interest

The authors have no conflict of interest.

References

1. Diochot S, Baron A, Salinas M, Douguet D, Scarzello S, Dabert-Gay A-S, et al. Black mamba venom peptides target acid-sensing ion channels to abolish pain. Nature. 2012;490(7421):552-5.
2. Shamloul M, Trusa J, Mett V, Yusibov V. Optimization and utilization of Agrobacterium-mediated transient protein production in Nicotiana. JoVE (Journal of Visualized Experiments). 2014(86):e51204.
3. Diochot S, Alloui A, Rodrigues P, Dauvois M, Friend V, Aissouni Y, et al. Analgesic effects of mambalgin peptide inhibitors of acid-sensing ion channels in inflammatory and neuropathic pain. Pain. 2016;157(3):552-9.

5.Schroeder CI, Rash LD, Vila-Farrés X, Rosengren KJ, Mobli M, King GF, et al. Chemical Synthesis, 3D Structure, and ASIC Binding Site of the Toxin Mambalgin-2. Angewandte Chemie International Edition. 2014;53(4):1017-20.
6.Boehm R. Bioproduction of therapeutic proteins in the 21st century and the role of plants and plant cells as production platforms. Annals of the New York Academy of Sciences. 2007;1102(1):121-34.
7.Lombardi R, Circelli P, Villani ME, Buriani G, Nardi L, Coppola V, et al. High-level HIV-1 Nef transient expression in Nicotiana benthamiana using the P19 gene silencing suppressor protein of Artichoke Mottled Crinckle Virus. BMC biotechnology. 2009;9(1):96.
8.Ma JK, Drake PM, Christou P. The production of recombinant pharmaceutical proteins in plants. Nature Reviews Genetics. 2003;4(10):794.
9.Porta C, Lomonossoff GP. Viruses as vectors for the expression of foreign sequences in plants. Biotechnology and Genetic Engineering Reviews. 2002;19(1):245-92.
10.Twyman RM, Stoger E, Schillberg S, Christou P, Fischer R. Molecular farming in plants: host systems and expression technology. TRENDS in Biotechnology. 2003;21(12):570-8.
11.Komarova TV, Baschieri S, Donini M, Marusic C, Benvenuto E, Dorokhov YL. Transient expression systems for plant-derived biopharmaceuticals. Expert Review of Vaccines. 2010;9(8):859-76.
12.Naseri Z, Khezri G, Davarpanah SJ, Ofoghi H. Virus-based vectors: a new approach for the production of recombinant proteins. Journal of Applied Biotechnology Reports. 2019;6(1):6-14.
13.Khezri G, Baghban Kohneh Rouz B, Ofoghi H, Davarpanah SJ. Heterologous expression of biologically active Mambalgin-1 peptide as a new potential anticancer, using a PVX-based viral vector in Nicotiana benthamiana. Plant Cell, Tissue and Organ Culture (PCTOC). 2020;142(2):241-51.
14.Naseri Z, Dorani Uliaei E, Ofoghi H, Davarpanah SJ. Expression of novel lhmlt fusion protein using plant viral vector and study of its anticancer effect. Plant Cell, Tissue and Organ Culture (PCTOC). 2020;143:97-108.
15.Menassa R, Ahmad A, Joensuu JJ. Transient expression using agroinfiltration and its applications in molecular farming. Molecular farming in plants: recent advances and future prospects: Springer; 2012. p. 183-98.
16.Salazar-González JA, Rosales-Mendoza S, Bañuelos-Hernández B. Viral vector-based expression strategies. Genetically Engineered Plants as a Source of Vaccines against Wide Spread Diseases: Springer; 2014. p. 43-60.
17.Ma Z, Zhu L, Song T, Wang Y, Zhang Q, Xia Y, et al. A paralogous decoy protects Phytophthora sojae apoplastic effector PsXEG1 from a host inhibitor. Science. 2017;355(6326):710-4.
18.Armenteros JJA, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, et al. SignalP 5.0 improves signal peptide predictions using deep neural networks. Nature biotechnology. 2019;37(4):420.
19.Jung S-K, McDonald K. Visual gene developer: a fully programmable bioinformatics software for synthetic gene optimization. Bmc Bioinformatics. 2011;12(1):340.
20.Sharp PM, Li WH. An evolutionary perspective on synonymous codon usage in unicellular organisms. J Mol Evol. 1986;24.
21.Sharp PM, Li WH. Codon usage in regulatory genes in Escherichia coli does not reflect selection for 'rare' codons. Nucleic Acids Res. 1986;14.
22.Rychlik W. OLIGO primer analysis software version 7. Molecular Biology Insights, Inc, Cascade, CO. 2008.
23.Brendel V, Kleffe J, Carle-Urioste JC, Walbot V. Prediction of splice sites in plant pre-mRNA from sequence properties. Journal of molecular biology. 1998;276(1):85-104.
24.Narsai R, Howell KA, Millar AH, O'Toole N, Small I, Whelan J. Genome-wide analysis of mRNA decay rates and their determinants in Arabidopsis thaliana. The Plant Cell. 2007;19 (11):3418-36.
25.Leite A, de Freitas FA, Yunes JA, Arruda P. Nucleotide sequence of a cDNA clone encoding γ -coixin from Coix lacryma-jobi seeds. Plant physiology. 1991;97(4):1604.
26.Jiang M-C, Hu C-C, Lin N-S, Hsu Y-H. Production of Human IFNγ Protein in Nicotiana benthamiana Plant through an Enhanced Expression System Based on Bamboo mosaic Virus. Viruses. 2019;11(6):509.
27.Carton JM, Sauerwald T, Hawley-Nelson P, Morse B, Peffer N, Beck H, et al. Codon engineering for improved antibody expression in mammalian cells. Protein expression and purification. 2007;55(2):279-86.
28.Webster GR, Teh AYH, Ma JKC. Synthetic gene design—The rationale for codon optimization and implications for molecular pharming in plants. Biotechnology and bioengineering. 2017;114(3):492-502.
29.Love AJ, Chapman SN, Matic S, Noris E, Lomonossoff GP, Taliansky M. In planta production of a candidate vaccine against bovine papillomavirus type 1. Planta. 2012;236(4):1305-13.
30.Perlak FJ, Fuchs RL, Dean DA, McPherson SL, Fischhoff DA. Modification of the coding sequence enhances plant expression of insect control protein genes. Proceedings of the National