ORIGINAL RESEARCH PAPER

Expression and characteristic of synthetic human epidermal growth factor (hEGF) in transgenic tobacco plants

Jie-Ying Bai · Lin Zeng · Yuan-Lei Hu · Yan-Fang Li · Zhong-Ping Lin · Shi-Chen Shang · Yan-Sheng Shi

Abstract To improve the accumulation of recombinant human epidermal growth factor (hEGF) in transgenic tobacco, a highly effective vector was constructed and transformed via *Agrobacterium tumefaciens*. The hEGF content in transgenic tobacco was up to 0.3% of the total soluble protein. Using the Vero E6 cell expansion assay and the MTT method for cell proliferation, hEGF produced by transgenic tobacco significantly stimulated Vero E6 cell expansion and proliferation, the same as commercial hEGF products.

Keywords Biological activity · Human epidermal growth factor · Matrix attachment regions · Optimized codons · Tobacco

J.-Y. Bai · Y.-L. Hu (✉) · Z.-P. Lin
The National Laboratory of Protein Engineering and Plant Genetic Engineering, Peking University, Beijing 100871, P.R. China
e-mail: huyl@pku.edu.cn

J.-Y. Bai · L. Zeng · S.-C. Shang · Y.-S. Shi
Laboratory Animal Center of the Academy of Military Medical Science, Beijing 100071, P.R. China

Y.-F. Li
College of Plant Sciences, Jilin University, Changchun 130062, P.R. China

Introduction

Pharmaceutical proteins, such as antibodies (Conrad and Fiedler 1998) and vaccines (Mason and Arntzen 1995), can be produced in plant systems, which is advantageous because the process is easy and low-cost (Daniell et al. 2001). The proteins are free from pathogens (Ma and Hein 1996), the purification procedure is simple and it is also possible to produce edible therapeutic products. However, there are some factors limiting the use of plant-based production of pharmaceutical proteins, such as that limited expression level of the target genes might be too low for any practical application. Therefore, it is a key task to increase the accumulation of heterologous proteins in transgenic plants.

To gain accurate expression of foreign protein in large quantities, genetic modification of the target protein is often needed, such as strong promoters and correct signal sequences for protein sub-cellular localization. Also the choice of the recipient plant may be critical. The utilization of a four-amino-acid endoplasmic reticulum retention signal peptide, Lys-Asp-Glu-Leu (KDEL), has led to the accumulation of several proteins in the endoplasmic reticulum and protects them from being destroyed by enzymes in the cytoplasm. Matrix attachment regions (MARs) could specifically bind the nuclear matrix and form chromatin loops in eukaryotic cells. The use of MARs would result in alleviated homologous gene silencing and boost stable protein accumulation of exogenous genes (Strick and Laemmli 1995; Lee et al. 1998).

Higo et al. (1993) tried to express human epidermal growth factor (hEGF) in tobacco plants, but the recombinant protein harvested only accounted for 0.001% of the total soluble proteins. Wirth et al. (2004) constructed two different vectors that caused the recombinant protein to be targeted into either the cytoplasm or the apoplast. Recombinant hEGF with biological activity was accumulated in the endoplasmic reticulum at 0.1% of the total soluble proteins. In this work, the following strategies were adopted to elevate hEGF expression in transgenic tobacco: (1) using plant bias codon in the reading frame, (2) adding the endopalsmic reticulum retention signal peptide KDEL at the C-terminal of the protein, and (3) harboring MARs at both ends of the hEGF transcriptional unit. Our results showed that the recombinant hEGF accumulated in transgenic tobacco plants amounted up to 0.3% of the total soluble proteins. Meanwhile, recombinant hEGF presented similar biological activity as commercialized hEGF, according to cell expansion and proliferation assays.

Material and methods

Design of the human epidermal growth factor gene

The nucleotide sequence of hEGF was modified according to the codon bias of tobacco plants (Willbur and Gowri 1990, Table 1), and a sequence coding the four-amino-acid endoplasmic reticulum retention signal peptide KDEL, AAGGATGAGCTC, was added at the 3′-terminal. The designed hEGF protein contains a 177 bp open reading frame and a KDEL signal at its C-terminal, and the amino acid sequence is as the following: MNSDSECPLSHDGY CLHDGVCMYIEALDKYACNCVVGYIGERCQY RDLKWWELRKDEL. The synthetic hEGF gene was amplified by assembly PCR using a set of overlapping 20-mer oligonucleotides with Taq DNA polymerase.

Plant transformation

The binary expression vectors pCAMhEGF, pCAMhEGFM and pCAMBIA3300 were transferred into A. tumefaciens strain LBA4404 by the freeze-thaw method. The leaf discs of Nicotiana tabacum L. var.

Table 1 Comparison of codon usage in wild-type and synthetic human epidermal growth factor (hEGF)

| Amino acid	Codon of hEGF	
	Wild type	Synthetic
Asn	AAT	AAC
Ser	TCC TCT$_2$	AGC TCC TCT
Asp	GAC GAT$_4$	GAT$_5$
Glu	GAA$_4$	GAG$_4$
Cys	TGT$_4$ TGC$_2$	TGC$_6$
Pro	CCG	CCA
Leu	CTG$_5$	CTT$_4$ CTC
His	CAT CAC	CAC$_2$
Gly	GGT$_4$	GGA$_3$ GGT
Val	GTT$_2$ GTA	GTT$_3$
Ala	GCA$_2$	GCT$_2$
Lys	AAA$_2$	AAG$_2$
Trp	TGG$_2$	TGG TAG

Codons of the synthetic gene were decided according to the codon bias of tobacco, host plant of the transgene in this study. The footnote number of the codons represents the times of their appearance in the full-length sequence.

W38 were co-cultured with A. tumefaciens and plantlets differentiated on MS medium supplemented with 1 mg benzyl adenine (6-BA)/L, 5 mg phosphinotricin (PPT)/L and 500 mg carbenicillin (Cb)/L. The regenerated shoots (2–3 cm) were transferred into hormone-free 1/2 MS medium with 5 mg PPT/L and 500 mg Cb/L for rooting. Well-rooted plants were subsequently transferred into soil and kept under greenhouse conditions.

Results

Expression of human epidermal growth factor in transgenic tobacco

pCAMhEGFM and pCAMhEGF, two genetic constructs for plant transformation (Fig. 1), were introduced into tobacco via A. tumefaciens. The transformants were screened with Northern blot (Fig. 2). The highest hEGF expression levels in pCAMhEGF tobacco was only 0.09% of the total soluble protein, while in pCAMhEGFM tobacco the highest level was up to 0.3% of the total soluble protein. As shown in Fig. 3, the average amounts of recombinant hEGF in leaves were 0.03 mg/g FW in
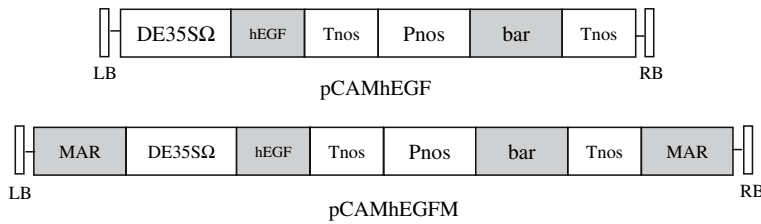
Fig. 1 Constructs of binary vectors pCAMhEGF and pCAMhEGFM for tobacco transformation. The synthetic human epidermal growth factor (hEGF) was subcloned into pCAMBIA3300 (CAMBIA) to form the expression cassette pCAMhEGF, in which the gene was under the control of a duplicated CaMV35S promoter with Ω enhancer (DE35SΩ). The DE35SΩ-hEGF-Tnos frame was further subcloned between two matrix attachment regions (MARs) to form the expression cassette pCAMhEGFM. These two vectors also had bar expression cassette as the selection marker against phosphinotricin in transgenic tobacco. LB and RB, left border and right border of binary vectors; Pnos and Tnos, promoter and terminator of the *nos* gene from *A. tumefaciens*

pCAMhEGF tobacco and 0.31 mg/g FW in pCAM-hEGFM tobacco, respectively.

Characterization of recombinant human epidermal growth factor

The biological activity of recombinant hEGF was detected by cell proliferation assays using the MTT method (Mosmann 1983). As shown in Fig. 4, the cell cultures treated with protein extracts from transgenic plants exhibited more prominent color change than those treated with protein extracts from w38, indicating a more active cell mitosis in the former ones due to the mitogenic activity of recombinant hEGF; besides, when the same quantities of proteins were supplemented, the cell culture supplemented with extracts from pCAMhEGFM tobacco presented a deeper color, indicating a higher hEGF content, than those supplemented with extracts from pCAMhEGF tobacco. A_{492} of the cell cultures was measured and analyzed using the Chi square test (Table 2), which showed significant differences ($P < 0.001$) among wild-type tobacco, pCAMhEGFM transformants and pCAMhEGF transformants.
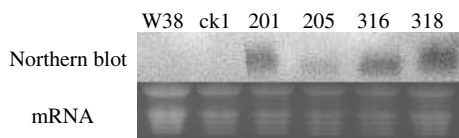
Fig. 2 Northern blot analysis of transgenic tobacco lines. Total RNA was isolated from 0.5 g tobacco leaves by TRIzol Reagent (Invitrogen). RNA (about 10 μg) was separated on 1% denaturing agarose gel containing 1% (w/v) formaldehyde, blotted onto Hybond N+ nylon membranes (Millipore) and then fixed by UV cross-linking. Prehybridization and hybridization of the membrane were performed at 68°C. Specific mRNA sequences were detected by hybridization with digoxigenin-dUTP labeled human epidermal growth factor cDNA probes following instructions of the DIG DNA Labeling and Detection Kit (Catalog No.1093657, Boehringer Mannheim). W38, wild-type plant; Ck1, tobacco transformed with pCAMBIA3300; 201 and 205, two tobacco lines transformed with pCAMhEGF; 316 and 318, two tobacco lines transformed with pCAMhEGFM. Ethidium bromide dyed RNA gel was shown as a comparison
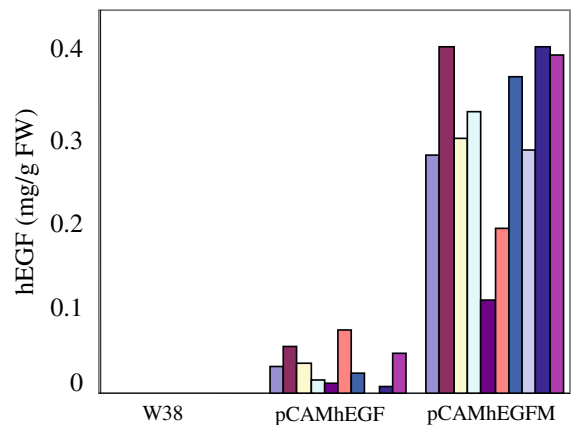
Fig. 3 hEGF accumulation in transgenic plants. hEGF expressed in transgenic plants was quantified in leaf tissue extracts by enzyme-linked immunoassay (ELISA) using commercial hEGF protein (PeproTech) as the standard. The monoclonal rabbit anti-hEGF antibody (1:1000, 100 ng/ml, PeproTech) and the alkaline phosphatase-linked goat anti-rabbit antibody (1:2000, Sigma) were employed respectively as the first and the second antibody interacting in ELISA. Reaction was stopped by adding 2 M H_2SO_4 and hEGF content was determined based on A_{492}. Genetic constructs used for transformation are indicated under each group. W38: wild-type control
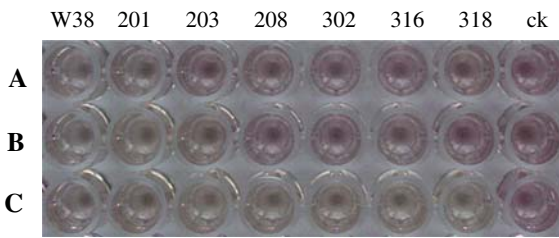
Fig. 4 Mitogenic activity of tobacco-expressed hEGF on Vero E6 cells. Vero E6 cells, treated with trypsin, were seeded in 96-well culture plates at 10^5 cells/ml and 100 μl/well. A serial dilution of commercial hEGF and tobacco-expressed total proteins were supplemented into the wells. After 8 h incubation, MTT was added into the wells, and the cultures were incubated for another 4 h. During this period, mitochondrial enzymes will catalyze MTT into formazan (Mosmann 1983), which is deep purple. A_{492} of the cell culture can be measured using a 96-well plate ELISA reader for quantification. W38, supplemented with extracts from wild-type tobacco; 201, 203 and 208, supplemented with extracts from 3 individual pCAMhEGF transgenic tobacco lines; 302, 316 and 318, supplemented with extracts from 3 individual pCAMhEGFM transgenic tobacco lines; ck, supplemented with commercial hEGF. A, B and C, supplemented with 100 μg/ml, 50 μg/ml and 10 μg/ml total solution protein from transgenic tobacco, or 1 μg/ml, 0.5 μg/ml and 0.1 μg/ml commercial hEGF, respectively

The biological properties of the hEGF were further characterized through cumulus expansion assays. Dialyzed extracts from pCAMhEGFM tobacco line 316 and pCAMhEGF tobacco line 201 had the similar activity to promote cell expansion as commercial hEGF, as shown in Fig. 5A, B and C, while no cell expansion was observed in control (Fig. 5D).

Discussion

It is helpful to optimize the codons of animal or microorganism genes before expressing them in plants. Codon-optimized *Bacillus thuringiensis* (Bt) crystal insecticidal protein gene, which is currently widely-used in plant anti-insect bioengineering has offered us an ideal example (Perlak et al. 1991). With the identical amino acid sequence as the original hEGF, the synthetic gene had an increased GC content of 49.7%, compared with 45.5% in the wild-type protein. As plant genome generally has a higher GC component (Perlak et al. 1991), increase in GC content will likely be beneficial for the expression of exogenous genes with low GC content in plants. As we chose tobacco as the transformation host, designing a new gene with tobacco-biased codons was believed to contribute to protein expression level.

Targeted expression of the exogenous protein is another effective strategy to increase its yield. Wirth et al. (2004) once expressed recombinant hEGF (non-optimized codons) in tobacco and got a highest expression level of 0.11% of the total soluble proteins, which was 10 times higher than that obtained by Higo et al. (1993). He reported that when hEGF was targeted to apoplast, the expression level was 10^4 higher than expressing it in cytoplasm, indicating that an integrative system might be vital for hEGF expression in transgenic plants. There have been numerous studies focusing on producing industrial and pharmaceutical proteins in targeted cells or sub-cellular organelles, such as production of avidin in maize seeds for commercial exploitation (Hood et al. 1997), and highly expression of a secreted protein, human somatotropin (hST) in chloroplasts (Jeffrey et al. 2000). In animal and plant cells, KDEL and HDEL are two signal peptides playing important roles in retention of expressed proteins in the endoplasmic reticulum (Munro and Pelham 1987; Pelham et al. 1988; Napier et al. 1992), so we used KDEL to target the recombinant protein in order to get an elevated accumulation.

Table 2 Chi square test result of A_{492} of cell cultures in the cell proliferation assay

Difference source	Sum of squares	Degrees of freedom	Mean squares	*F*-ratio	*P*-value[b]	Fcrit
Between the teams[a]	0.0049	3	0.0016	13.3	5.34E-05	3.1
In the team	0.0025	20	0.0001			
Total	0.0074	23				

[a] The three teams in this test were wild-type, pCAMhEGF and pCAMhEGFM tobacco

[b] Probability according to *F*-test for homogeneity of variance compared with the three teams

A *P*-value under 0.001 means significant differences among the teams. For a detailed protocol of the cell proliferation assay, please refer to the legend of Fig. 4
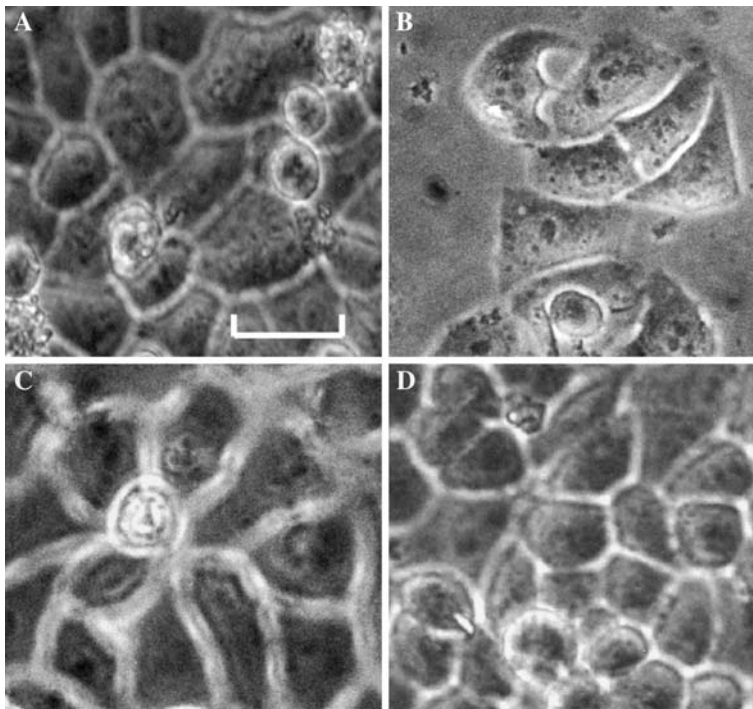
Fig. 5 Vero E6 cells cumulus expansion tests. Vero E6 cells were cultivated in cell culture medium RPMI-1640 (Invitrogen) supplemented with 10% fetal bovine serum (FBS, W/V), 100 units penicillin G/ml, and 100 µg streptomycin/ml on 96-well culture plates at a density of 10^4 cells/well. The plates were kept at 37°C in a humidified atmosphere of 5% CO_2 (v/v) in the air for 4 h. Afterwards, extracts containing 10 ng hEGF from pCAMhEGF line 201 (**A**) or pCAMhEGFM line 316 (**B**), 10 ng commercial hEGF (**C**), or extracts from wild-type tobacco (**D**) was supplemented into the wells, and the plates were cultured for another 8 h at 37°C with 5% CO_2 (v/v) in the air. Cumulus expansion was visually assessed under a stereomicroscope. Scale bar: 50 µm

MARs protect foreign gene from being silenced both in transgenic plants and animals (Zhong et al. 2004; Lee et al. 1998; Whitelaw et al. 2000). The MARs employed in our study was isolated from yeast whose function of enhancing the expression of foreign proteins has been well established (Zhong et al. 2004). Similarly, it showed a prominent effect in escorting hEGF heterogenous expression, resulted in remarkable increase in protein yield (Fig. 3).

The highest hEGF level in one of the transgenic tobacco lines reached 0.3% of the total soluble proteins, which was the highest ever documented, indicating that the above explorations, such as optimized codons, sub-cellular targeting and MARs employment were profitable in hEGF overexpression in transgenic tobacco. Future studies are contemplated to focus on further improving hEGF expression, and expressing the protein in other plant hosts.

References

Conrad U, Fiedler U (1998) Compartment-specific accumulation of recombinant immunoglobulins in plant cells: an essential tool for antibody production and immunomodulation of physiological functions and pathogen activity. Plant Mol Biol 38:101–109

Daniell H, Streatfield SJ, Wycoff K (2001) Medical molecular farming: production of antibodies biopharmaceuticals and edible vaccines in plants. Trends Plant Sci 6:219–226

Higo K, Saito Y, Higo H (1993) Expression of a chemically synthesized gene for human epidermal growth factor under the control of mosaic virus 35S Promoter in Transgenic Tobacco. Biosci Biotechnol Biochem 57:1477–1481

Hood EE, Witcher DR, Maddok S, Meyer T, Baszczynski C, Bailey M, Flynn P, Register J, Marshall L, Bond D, Kulisek E, Kusnadi A, Evangelista R, Nikolov Z, Wooge C, Mehigh RJ, Hernan R, Kappel WK, Ritland D, Ping LC, Howard JA (1997) Commercial production of avidin from transgenic maize: characterization of transformant, production, processing, extraction and purification. Mol Breed 3:291–306

Lee TH, Kim SJ, Han YM, Yu DY, Lee CS, Choi YJ, Moon HB, Baik MG, Le KK (1998) Matrix attachment region sequences enhanced the expression frequency of a whey acidic protein/hman lactoferrin fusion gene in the mammary gland of transgenic mice. Mol Cells 8:530–536

Ma JK, Hein MB (1996) Antibody production and engineering in plants. Ann NY Acad Sci 792:73–81

Mason HS, Arntzen CJ (1995) Transgenic plants as vaccine production systems. Trends Biotechnol 13:388–392

Mosmann T (1983) Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. J Immunol Methods 65:55–63

Munro S, Pelham HR (1987) A *C*-terminal signal prevents secretion of luminal ER proteins. Cell 48:899–907

Napier RM, Fowke LC, Hawes C, Lewis M, Pelham HR (1992) Immunological evidence that plants use both HDEL and KDEL for targeting proteins to the endoplasmic reticulum. J Cell Sci 102:261–271

Pelham HR, Hardwick KG, Lewis MJ (1988) Sorting of soluble ER proteins in yeast. EMBO J 7:1757–1762

Perlak FJ, Fuchs RL, Dean DA, McPherson SL, Fischhoff DA (1991) Modification of the coding sequence enhances plant expression of insect control protein genes. Proc Natl Acad Sci USA 88:3324–3328

Staub JM, Garcia B, Graves J, Peter TJH, Priscilla H, Narender N, Vikram P, Michael S, James AC, Lori S, Dannette W, Guangning Y, Douglas AR (2000) High-yield production of a human therapeutic protein in tobacco chloroplasts. Nature Biotechnol 18:333–339

Strick R, Laemmli UK (1995) SARs are cis DNA elements of chromosome dynamics: synthesis of a SAR repressor protein. Cell 83:1137–1148

Whitelaw CB, Grolli S, Accornero P, Donofrio G, Farini E, Webster J (2000) Matrix attachment region regulates basal beta-lactoglobulin transgene expression. Gene 244:73–80

Willbur HC, Gowri G (1990) Codon usage in higher plants, green algae, and cyanobacteria. Plant Physiol 92:1–11

Wirth S, Calamante G, Mentaberry A, Bussmann L, Lattanzi M, Barañao L, Bravo-Almonacid F (2004) Expression of active human epidermal growth factor (hEGF) in tobacco plants by integrative and non-integrative systems. Mol Breed 13:23–35

Zhong J, Liu SJ, Ma SS, Yang W, Hu YL, Wu Q, Lin ZP (2004) Effect of matrix attachment regions on resveratrol production in tobacco with transgene of stilbene synthase from *Parthenocissus henryana*. Acta Bot Sin 46:948–954