

Short communication

**IMPROVED FUSION PROTEIN EXPRESSION OF EGFP VIA THE
MUTATION OF BOTH KOZAK AND THE INITIAL ATG CODON**

CHAO DAI[#], ZHIJIAN CAO[#], YINGLIANG WU, HONG YI, DAHE JIANG
and WENXIN LI*

State Key Laboratory of Virology, College of Life Sciences, Wuhan University,
Wuhan, 430072, P R China

Abstract: Since its discovery, green fluorescence protein (GFP) has been used as a reporter in a broad range of applications, including the determination of gene expression in diverse organisms, and subcellular protein localization. pEGFP-N1 is a eukaryotic expression vector encoding EGFP, the MCS of which locates at the N terminus of EGFP. In this study, the cDNA sequence of scorpion toxin BmKK2 was inserted into the XhoI-HindIII cut of pEGFP-N1 to construct a toxin-EGFP fusion gene (named pEGFP-BmKK2). Fluorescence imaging revealed that HEK 293T cells that were transfected by pEGFP-BmKK2 emitted green fluorescence. Transcription of pEGFP-BmKK2 was confirmed by RT-PCR. However, western blotting analysis showed that the transfected HEK 293T cells expressed mostly EGFP, but little toxin-EGFP fusion protein, implying that pEGFP-N1 cannot be used as a fusion expression vector for subcellular protein localization for the BmKK2 gene. Consequently, two modified recombinant vectors (pEGFP-BmKK2-M1 and pEGFP-BmKK2-M2) were constructed based on pEGFP-BmKK2. This greatly improved the expression of toxin-EGFP fusion protein from pEGFP-BmKK2-M2.

[#] These authors contributed equally to this article

* Author for correspondence; e-mail: liwxlab@whu.edu.cn, tel: 86-27-68752831, fax: 86-27-68752146

Abbreviations used: CCTCC – China Center for Type Culture Collection; cDNA – complementary DNA; DAB – diaminobenzidine; DMEM – Dulbecco's modified Eagle's medium; *E. coli* – *Escherichia coli*; EGFP – enhanced green fluorescent protein; FCS – fetal calf serum; GFP – green fluorescence protein; MCS – multiple cloning site; mRNA – messenger RNA; RT-PCR – reverse transcript polymerase chain reaction; SDS-PAGE – sodium dodecyl sulfate polyacrylamide gel electrophoresis

Key words: EGFP, Fusion protein expression, Subcellular protein localization, Scorpion toxin

INTRODUCTION

Green fluorescence protein (GFP) was discovered in the jellyfish *Aequorea victoria*, and has been widely used as a reporter in studies of eukaryotic cells since the 1990s' [1]. Heterologous expression of GFP generates striking green fluorescence in many different organisms without additional substrates or cofactors [1], making it an excellent reporter for *in vivo* use.

The eukaryotic expression vector pEGFP-N1 has been widely used as a fusion protein vector for many years [2-6]. In most cases, pEGFP-N1 is used for subcellular localization by fusion protein expression. Here, to study the effect of toxin genes on cells, the cDNA sequence of BmKK2 (a short-chain potassium-channel scorpion-toxin gene) [7, 8] was inserted into the *XhoI-HindIII* cut of pEGFP-N1 to form a toxin-EGFP fusion gene. The constructed recombinant was named pEGFP-BmKK2. However, after HEK 293T cells were transfected with the toxin-EGFP fusion protein expression vector (pEGFP-BmKK2), western blotting analysis showed that there was almost no toxin-EGFP fusion protein expressed in the sample. For the purpose of improving fusion protein expression, pEGFP-BmKK2 was modified twice, either by the insertion of the Kozak sequence before the toxin-EGFP fusion gene (named pEGFP-BmKK2-M1) or by the insertion of the Kozak sequence as above and the mutation of the initial ATG codon of the EGFP gene (named pEGFP-BmKK2-M2). Western blotting analysis showed that toxin-EGFP expression was greatly improved in the HEK 293T cells that were transfected with the modified pEGFP-BmKK2-M2 recombinant.

MATERIALS AND METHODS

Construction of recombinant vectors

All the recombinant vectors used in this study were prepared using standard cloning techniques. The cDNA sequence of BmKK2 was obtained from Zeng *et al.*'s paper [9], and was amplified from the cDNA library constructed in our laboratory for the scorpion *Buthus martensii* Karsch, with the forward primer FP1 (5'-CGCCTCGAGATGAAAATATTTTTTGCTATTCTGCTCATA-3', corresponding to the initial translated region of BmKK2), and the reverse primer RP1 (5'-CGCAAGCTTTGTACAAGCGCAAATCCATTTC-3', corresponding to the terminal translated region of BmKK2 without the terminal codon for forming the toxin-EGFP fusion gene). The cDNA was inserted into the *XhoI-HindIII* cut of pEGFP-N1, yielding the recombinant vector pEGFP-BmKK2. At the same time, the recombinant pEGFP-BmKK2 was modified by adding the Kozak sequence (GCCACC) before the BmKK2-EGFP fusion gene and using the forward primer FP2 (5'-GGGCTCGAGGCCACCATGAAAATATTTTTTGC-3'). The thus-

modified pEGFP-BmKK2 was named pEGFP-BmKK2-M1. Forward primer FP3 (5'-CGGGATCCACCGGTCGCAACCTTGGTGAGC-3', corresponding to the initial translated region of EGFP) and reverse primer RP3 (5'-GTCGCGGCCGCTTACTTGTAC-3', corresponding to the terminal translated region of EGFP) were designed and synthesized to mutate the Kozak sequence and the initial ATG codon of EGFP in the recombinant pEGFP-BmKK2-M1. The thus-modified pEGFP-BmKK2-M1 was named pEGFP-BmKK2-M2. All the primers and recombinant vectors mentioned can be found in Fig. 1 or 2. All the recombinants were transformed to *E. coli* DH5 α , and then positive clones were screened and confirmed by sequencing.

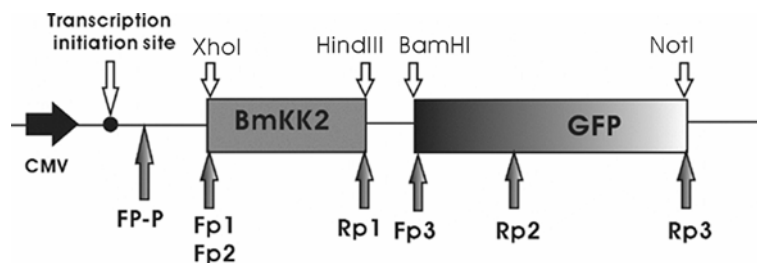


Fig. 1. Diagram showing the location of the primers (FP1, FP2, FP3, RP1, RP2, RP3 and FP-P) in the toxin-EGFP fusion gene of the recombinant vector. Primers FP1, FP2, FP3, RP1 and RP3 were used to construct the recombinants. Primers FP-P and RP2 were used for RT-PCR analysis.

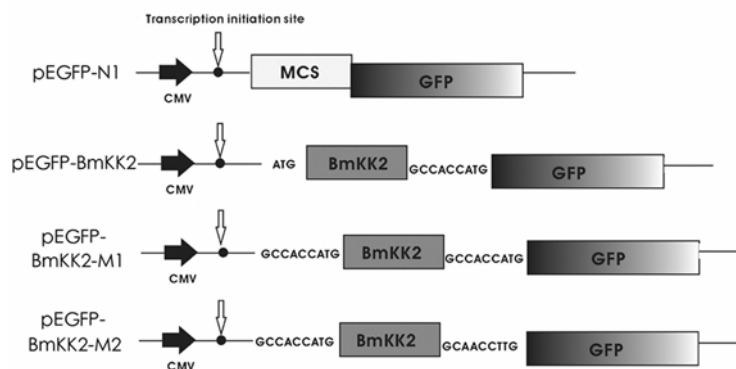


Fig. 2. Sketch maps of pEGFP-N1 and the recombinants (pEGFP-BmKK2, pEGFP-BmKK2-M1 and pEGFP-BmKK2-M2).

Cell culture and transfection assays

The HEK 293T cells were purchased from the China Center for Type Culture Collection (CCTCC). Cell cultures were grown in Dulbecco's modified Eagle's medium (DMEM) with fetal calf serum (FCS). Cells were plated at 7×10^4 in

six-well plates for 18 to 24 hours before transfection. The transfections were carried out using 1 μ g of each recombinant vector complexed with Sofast transfection reagent (Sunma Technology, P.R.China) as per the manufacturer's protocols. The transfection efficiency of each plate could easily be investigated via fluorescent microscopy.

RT-PCR assays

Total RNA was extracted with Trizol reagent 36 hours after transfection. The reverse transcript (RT) for cDNA synthesis was performed as normal. The total RNA was completely digested by RNase-free DNase (Promega, USA) to avoid contamination of the genomic DNA and the transfected plasmid DNA. After incubation at 42°C as a template for RT, the digested RNA products were precipitated to remove salts of the RT reaction (e.g. MgCl₂, dNTPs). The purified RT products were used as a PCR template. PCR consisted of 30 cycles and was carried out using FP-P and RP2 primers. FP-P (5'-GTCAGATCCGCTAGCGCTACC-3'), corresponding to the sequence between the transcription and translation initial sites of the toxin-EGFP fusion gene) was used as forward primer, and the reverse primer was RP2 (5'-CTGGACGTAGCCTTCGGGCATG -3', pairing with FP-P, and corresponding to 262-284 nt of the GFP encoding region). The resulting PCR products were cloned into the pGEM-T vector and sequenced.

Western blotting analysis

Cells transfected with pEGFP-N1, pEGFP-BmKK2-M1 or pEGFP-BmKK2-M2 were suspended with micropipette tips 48 hours after transfection. The cells were boiled in SDS loading buffer, and then run on 12% SDS-PAGE. Transmembrane blotting was done at 25 V overnight. After washing in PBS for 10 min, the membrane was incubated in blocking solution (5% Skim milk) for 30 min and then incubated with rabbit anti-GFP antibody (1:1000, Sanying, Technology, P.R.China) for 3 hours. After washing several times for 20 minutes, the membrane was incubated with goat anti-rabbit IgG (1:1000, Sanying, Technology, P.R.China) for 1 hour. The membrane was stained with diaminobenzidine reagent (DAB, Sigma). All the experiments were repeated.

RESULTS AND DISCUSSION

Fluorescence microscopy revealed that 24 hours after transfection, the fluorescence intensity of the sample cells transfected with pEGFP-N1 vector was the strongest (Fig. 3). The data implied that the total amount of toxin-EGFP fusion protein expressed from the other recombinant samples was less than the amount of protein from the cells transfected with pEGFP-N1.

RT-PCR was carried out to ascertain the expression of toxin-EGFP fusion protein in the transfected HEK 293T cells at the mRNA level. PCR by untranscribed RNA, obtained from any of the four cell samples as a template, with FP-P and RP2 as the sense and antisense primers, did not amplify the

products. These results revealed that RNA samples from the transfected HEK 293T cells were not contaminated by the transfected recombinant plasmids (Fig. 4, lanes 1, 3, 5 and 7). RT-PCR with FP-P and RP2 as sense and antisense primers was used to detect the expression of toxin-EGFP fusion gene mRNA. RT-PCR for the pEGFP-N1 sample amplified only one fragment at the same length as the PCR products for the pEGFP-N1 plasmid as the template. The RT-

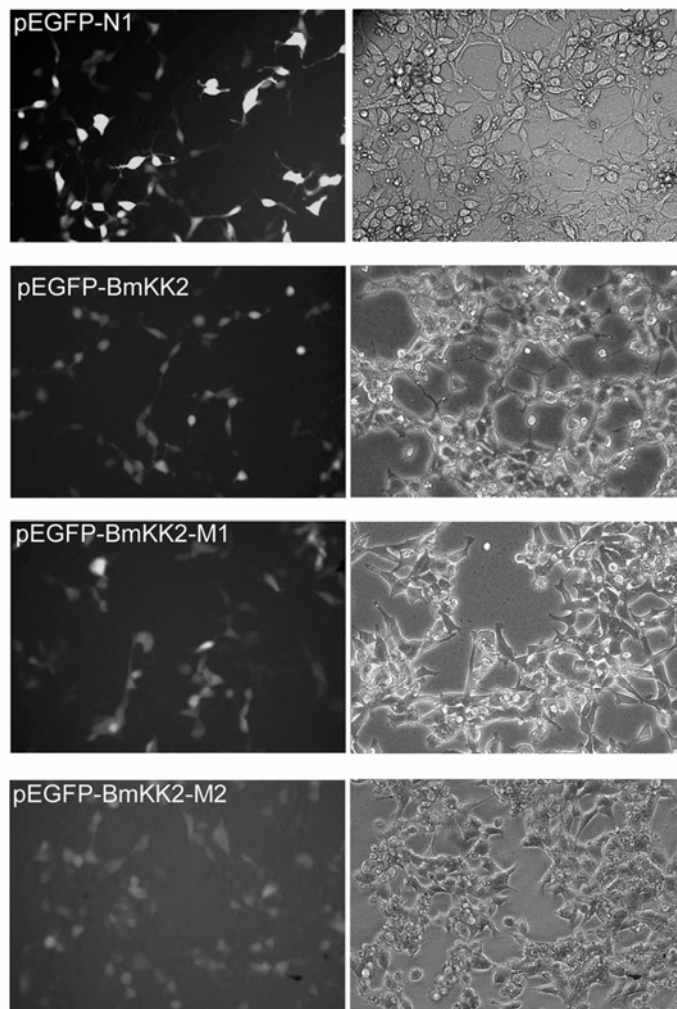


Fig. 3. Forty eight hours after transfection, the cells were imaged via fluorescence microscopy. This figure shows the GFP and toxin-EGFP fusion protein expression in the HEK 293T cells transfected with pEGFP-N1, pEGFP-BmKK2, pEGFP-BmKK2-M1 and pEGFP-BmKK2-M2, respectively. The left panels show the EGFP and toxin-EGFP fusion protein expression visualized under blue fluorescence, while the right panels showed the corresponding cells visualized under bright light.

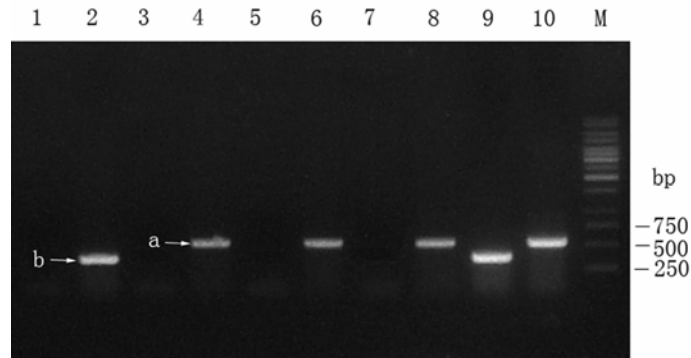


Fig. 4. Forty eight hours after transfection, the RT-PCR of the toxin-EGFP gene expression at the mRNA level. Lanes 1, 3, 5 and 7: RT-PCR products with four cell samples' RNase-free Dnase digested RNAs as template (cell samples transfected with pEGFP-N1, pEGFP-BmKK2, pEGFP-BmKK2-M1 or pEGFP-BmKK2-M2), and FP-P and RP2 as the sense and antisense primers. Lanes 2, 4, 6 and 8: RT-PCR products with four cell samples' RT as template (cell samples transfected with pEGFP-N1, pEGFP-BmKK2, pEGFP-BmKK2-M1 or pEGFP-BmKK2-M2), and FP-P and RP2 as the sense and antisense primers. Lanes 9 and 10: PCR products of positive controls with pEGFP-N1 and pEGFP-BmKK2 recombinants as templates, and FP-P and RP2 as the sense and antisense primers. M: DNA marker, 1-kb ladder. The band marked a is the same size as lane 10, and b is the same size as lane 9, which proved that all of the samples were well transcribed.

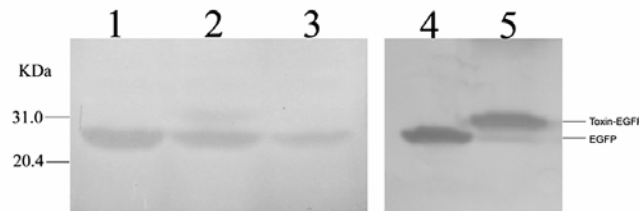


Fig. 5. Forty eight hours after transfection, EGFP and toxin-EGFP fusion protein expression analysis by western blotting. Each recombinant vector was transfected into HEK 293T cells, and the total cell lysates were used as samples of western blotting analysis. Left panel: the fusion protein expression was partly improved as in the recombinant vector pEGFP-BmKK2-M1. Right panel: after the Kozak sequence and initial ATG codon of EGFP gene were both mutated in pEGFP-BmKK2-M1, the toxin-EGFP fusion protein was successfully expressed. (1: pEGFP-N1; 2: pEGFP-BmKK2-M1; 3: pEGFP-BmKK2; 4: pEGFP-N1; 5: pEGFP-BmKK2-M2)

PCR products using pEGFP-BmKK2, pEGFP-BmKK2-M1 and pEGFP-BmKK2-M2 as templates were the same size as those obtained using pEGFP-BmKK2 plasmid as the template. These PCR results indicated that all of the transfected plasmids were properly transcribed.

The HEK 293T cells that were transfected with the recombinant vector pEGFP-BmKK2 expressed almost no toxin-EGFP fusion protein, whereas the cells that were transfected with the recombinant vector pEGFP-BmKK2-M1 expressed more toxin-EGFP fusion protein than the samples of pEGFP-BmKK2 (Fig. 5). We found that the proportion of the fusion protein improved a little with this method. However, there was still far more EGFP expressed than the toxin-EGFP fusion protein. Another modified recombinant vector, pEGFP-BmKK2-M2, was designed. It was different from pEGFP-BmKK2-M1 in that the Kozak sequence and the initial ATG codon of the GFP gene were both mutated (Fig. 2). As expected, the yield of toxin-EGFP fusion protein from the sample with pEGFP-BmKK2-M2 was greatly improved.

In this investigation, three recombinant plasmids (pEGFP-BmKK2, pEGFP-BmKK2-M1 and pEGFP-BmKK2-M2) were cloned and sequenced before transfection. As the sequencing primer is a universal primer of the pEGFP-N1 vector, cross contamination was avoided. On the other hand, the amounts of all the plasmids used for transfection were the same, and the cells prepared for transfection were from the same cell culture, so the effect on transfection efficiency can be ignored. After transfection, the cells were observed under the microscope both after 48 hours and 72 hours. It was found that the toxin had almost no toxic effects on cells: there was no obvious reduction in the number of cells.

Before our study, many researchers used pEGFP-N1 as a fusion protein expression vector [2-6]. Our results indicated that the toxin-EGFP fusion protein was hardly detected when pEGFP-N1 was used as a fusion expression vector. After modifying the vector twice, the toxin-EGFP fusion protein was properly expressed. As a result, the toxin-EGFP fusion protein can be successfully obtained with the recombinant vector pEGFP-BmKK2-M2. Clearly, the modified pEGFP-N1 vector, which consisted of mutations in the Kozak sequence and the initial ATG codon of the EGFP gene, can also be used for subcellular localization.

RT-PCR results showed that the toxin-EGFP genes of the three samples (pEGFP-BmKK2, pEGFP-BmKK2-M1 and pEGFP-BmKK2-M2) were properly transcribed, and that each product appeared to be one single mRNA. Although the ratio of the amount of EGFP and toxin-EGFP fusion protein differed between them, it was confirmed by western blotting analysis that both EGFP and the toxin-EGFP fusion protein were expressed from these samples. The results of RT-PCR and western blotting possibly suggested that both EGFP and toxin-EGFP fusion protein were translated from the toxin-EGFP fusion mRNA in pEGFP-BmKK2, pEGFP-BmKK2-M1 and pEGFP-BmKK2-M2. Some researchers have reported that a single mRNA encodes two or more independent proteins in eukaryotes [10, 11], which is an important molecular mechanism of the proteome complexity. We presume that this molecular mechanism is responsible for the expression of both EGFP and toxin-EGFP fusion protein from the samples mentioned in this article.

Acknowledgements. This study was supported by the grants from the National Natural Sciences Foundation of China to Li Wenxin, Wu YingLiang and Cao Zhijian (Numbers: 30530140; 30500089; 30570045) and the Provincial Natural Sciences Foundation of HuBei to Cao Zhijian (Number: 2005ABA116).

REFERENCES

1. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W. and Prasher, D.C. Green fluorescent protein as a marker for gene expression. **Science** 263 (1994) 802-805.
2. Gill, S.K., Bhattacharya, M., Ferguson, S.S. and Rylett, R.J. Identification of a novel nuclear localization signal common to 69- and 82-kDa human choline acetyltransferase. **J. Biol. Chem.** 278 (2003) 20217-20224.
3. Kamei, H. Intracellular localization of histidyl-tRNA synthetase/Jo-1 antigen in T24 cells and some other cells. **J. Autoimmun.** 22 (2004) 201-210.
4. Chen, T., Li, X., Yang, Y. and Church, R.L. Localization of lens intrinsic membrane protein MP19 and mutant protein MP19(To3) using fluorescent expression vectors. **Mol. Vis.** 8 (2002) 372-388.
5. Sun, D., Landowski, C.P., Chu, X., Wallsten, R., Komorowski, T.E., Fleisher, D. and Amidon, G.L. Drug inhibition of Gly-Sar uptake and hPepT1 localization using hPepT1-GFP fusion protein. **AAPS Pharm. Sci.** 3 (2001) E2.
6. Cormack, B.P., Valdivia, R.H. and Falkow, S. FACS-optimized mutants of the green fluorescent protein (GFP). **Gene** 173 (1996) 33-38.
7. Li, M.H., Zhang, N.X., Chen, X.Q., Wu, G., Wu, H., Hu, G.Y. Purification and pharmacological characterization of BmKK2 (alpha-KTx 14.2), a novel potassium channel-blocking peptide, from the venom of Asian scorpion *Buthus martensi* Karsch. **Toxicon** 43 (2004) 895-900
8. Rodriguez, D.A., Vega, R.C., Possani, L.D. Current views on scorpion toxins specific for K⁺-channels. **Toxicon** 43 (2004) 865-75.
9. Zeng, X.C., Peng, F., Luo, F., Zhu, S.Y., Liu, H. and Li, W.X. Molecular cloning and characterization of four scorpion K⁽⁺⁾-toxin-like peptides: a new subfamily of venom peptides (alpha-KTx14) and genomic analysis of a member. **Biochimie** 83 (2001) 883-889.
10. Gray, T.A., Saitoh, S. and Nicholls, R.D. An imprinted, mammalian bicistronic transcript encodes two independent proteins. **Proc. Natl. Acad. Sci. USA** 96 (1999) 5616-5621.
11. Sloan, J., Kinghorn, J.R. and Unkles, S.E. The two subunits of human molybdopterin synthase: evidence for a bicistronic messenger RNA with overlapping reading frames. **Nucleic Acids Res.** 27 (1999) 854-858.