The strategy of fusion genes construction determines efficient expression of introduced transcription factors

Tomasz Adamus, Paweł Konieczny, Małgorzata Sekuła, Maciej Sułkowski and Marcin Majka

Department of Transplantation, Jagiellonian University Medical College, Kraków, Poland

The main goal in gene therapy and biomedical research is an efficient transcription factors (TFs) delivery system. SNAIL, a zinc finger transcription factor, is strongly involved in tumor, what makes its signaling pathways an interesting research subject. The necessity of tracking activation of intracellular pathways has prompted fluorescent proteins usage as localization markers. Advanced molecular cloning techniques allow to generate fusion proteins from fluorescent markers and transcription factors. Depending on fusion strategy, the protein expression levels and nuclear transport ability are significantly different. The P2A self-cleavage motif through its cleavage ability allows two single proteins to be simultaneously expressed. The aim of this study was to compare two strategies for introducing a pair of genes using expression vector system. We have examined GFP and SNAI1 gene fusions by comprising common nucleotide polylinker (multiple cloning site) or P2A motif in between them, resulting in one fusion or two independent protein expressions respectively. In each case transgene expression levels and translation efficiency as well as nuclear localization of expressed protein have been analyzed. Our data showed that usage of P2A motif provides more effective nuclear transport of SNAIL transcription factor than conventional genes linker. At the same time the fluorescent marker spreads evenly in subcellular space.

Key words: multiple gene expression, P2A, self-cleavage motif, SNAIL, transcription factors introduction

INTRODUCTION

The ability to visualize, track and quantify biomolecules in cells is crucial for understanding biological mechanisms. Specific ‘point of interest’ in cell, like protein or nucleic acid can be fluorescently labeled by attaching a special compound called fluorophore. That compound has to meet several requirements such as high ability of visualization, minimal perturbation in examined target and appropriate indication of transfection efficiency. Proteins like green fluorescent protein (GFP) are commonly used as reporters of the expression. In case of transgene introduction, fluorescent proteins are considered to be the proof that given genes are expressed. They can also be used for examination of potential intracellular transport of introduced gene (Chudakov et al., 2010).

Introduction of transgene in the presence of reporter gene often requires strategies that allow co-expression of introduced genes, so several approaches have been employed. One of them is insertion of internal ribosomal entry sites (IRESs) between genes, which is widely used (Mokrejs, 2006). The other strategy is the usage of dual or multiple promoter systems, but it can be associated with differences in protein expression levels or silencing of some promoters (Radcliffe & Mitrophanous, 2004). It is also possible to introduce genes applying multiple vectors transfection, however uneven transcription efficiency has to be considered as well as increased toxicity (Arbab et al., 2004). The fusion genes under control of one promoter can be used as well, however expressed protein is an amino acid chain containing sequence of interest and reporter part, that can possibly affects the results.

The limitations of listed strategies can be overcome using 2A linker that belongs to the 2A peptides family. A ‘self-cleaving’ peptide has been at first identified in the foot-and-mouth disease virus (FMDV), a member of the picornavirus (Kim et al., 2011). The cleavage always takes place between glycine and proline at the C-terminus of the sequence (Szmyczak et al., 2004). It has been shown that creation of multiple proteins from one transcript is achieved at the translation step and the ‘ribosomal skip’ mechanism is postulated (Gao et al., 2012). Due to the fact that P2A motif is small in comparison to IRES sequence, it does not affect expression vector capacity, especially for multicistronic constructs, where couple of linkers have to be used. The P2A linker is potential target for the rapid development and testing of polycistronic constructs in the field of gene therapy strategies. The ability of feasible multiple genes introduction is desirable in biomedical research of cell reprogramming (Gao et al., 2012).

Efficient transcription factors (TFs) delivery system can be a main goal in gene therapies of cancer or other diseases and biomedical research (Guo et al., 2012; Yeh et al., 2013). One of TFs which seems to play a crucial role in tumor progression of numerous cancers is SNAI1, a zinc finger transcription factor (Lee et al., 2006). The target for TFs is nuclear DNA, thus the transport into nucleus is essential. Highly conserved motifs in SNAI1 gene sequence, known as nuclear localization signals (NLSs), responsible for its intracellular transport are present. SNAI1 transcription factor, like many others nuclear regulatory proteins, contains serine-rich domain

Abbreviations: FMDV, foot-and-mouth disease virus; IRESs, internal ribosomal entry sites; GFP, NLSs, nuclear localization signals; Proteins like green fluorescent protein; TFs, transcription factors

* e-mail: mmajka@cm-uj.krakow.pl
which is the phosphorylation site and can be phosphorylated by GSK3β kinase in cytoplasm. The result of the phosphorylation is β-TRCP-mediated ubiquitination and proteasomal degradation of the protein (Yook et al., 2005). Rapid proteolysis of SNAIL in cytoplasm suggests that proper nuclear active transport is essential for correct function of this TF (Yamasaki et al., 2005). Learning about SNAIL pathways may provide an experimental basis for the possible therapy strategies to decrease invasiveness of cancer cells (Batlle et al., 2000).

MATERIALS AND METHODS

Vectors generation. DNA constructs were generated by performing multiple PCR reactions introducing simultaneously restriction sites and recombination sequences (attB) at the ends of the inserts (Fig. 1). Enhanced green fluorescence protein (GFP) sequence was amplified with pEGFP-C1 plasmid (Clontech, USA), human SNAIL (SNAIL) with GFP SNAIL WT plasmid (Addgene, USA) and P2A with FUW-OSKM plasmid (Addgene) as templates. Each amplicon was digested with appropriate restriction enzyme and purified using GeneMATRIX DNA Purification Kit (EURx, Poland). Further, GFP digested fragment was dephosphorylated with calf Intestinal Alkaline Phosphatase (EURx) and ligated with SNAIL fragment using T4 DNA Ligase (EURx) to obtain GFP-SNAIL fusion construct, or subsequently ligated with P2A then SNAIL to get GFP-P2A-SNAIL. All the constructs were reamplified using attB flanking primers (Fig. 1C), respectively GFPattB forward (5′-GGGGACCACTTTTGATACAAAGAGGCTATCCAGTGTAGGAACGAGGCGA) and SNAILattB reverse (5′-GGGGACCACTTTTGATACAAAGAGGCTATCCAGTGTAGGAACGAGGCGA). Both GFP-SNAIL and GFP-P2A-SNAIL coding sequences started with Kozak consensus (ACCATG) at the 5′ of GFP and ended with stop codon (ATG) at the 3′ of SNAIL. Purified products were directly recombined with pDONR221 vector following recombination with expression vector using BP and LR clonase, respectively (Invitrogen, USA). Obtained clones were verified by restriction analysis and sequencing as well. As a control vector GFP@pLenti6/V5 was used, previously developed in our laboratory.

Cell culture and transfection. HEK293T human embryonic kidney and HeLa human cervical cell lines (ATCC, USA) were cultured respectively in DMEM (PAA, Austria) or MEM (PAA) supplemented with 10% FBS (PAA) and human SNAIL (SNAIL) gene (Hs00195591_m1; Applied Biosystems) and human vimentin (VIM) gene (Hs00185584_m1; Applied Biosystems) were amplified by TaqMAN Gene Expression Master Mix (Applied Biosystems) and primer-probe for GFP (MR04329676_m1; Applied Biosystems), human SNAIL (SNAIL) gene (Hs00195591_m1; Applied Biosystems) and human vimentin (VIM) gene (Hs00185584_m1; Applied Biosystems). The mRNA expression levels were normalized to the human housekeeping gene GAPDH (4352934; Applied Biosystems) while ddCt method (2−∆∆Ct) was used to calculate relative expression of the genes.

Western Blot analysis. HEK293T cells were collected 48h post transfection and lysed (for 10 min) on ice in M-PER buffer ( Pierce Rockford, USA) containing protease and phosphatase inhibitors (Sigma) to obtain total protein fraction, and the nuclear fraction was isolated using NEB-B buffer (29 mM Hepes, pH 7.9; 0.4 M NaCl; 1 mM EDTA; 1 mM EGTA). The protein content was determined by Bradford method. Subsequently, the extracted proteins were separated on 12% sodium dodecyl sulphate polyacrylamide electrophoresis gel (SDSPAGE), then fractioned proteins were transferred onto a PVDF membrane (BioRad). After blocking in 1% BSA TTBS buffer (Twee Tris Buffered Saline) membranes were incubated overnight with primary rabbit polyclonal antibodies for SNAIL protein (1:200; sc-28199; Santa Cruz Biotechnology, USA) as well as for GFP protein (1:1000; 2555; Cell Signaling Technology, USA), respectively, and subsequently detected with horseradish peroxidase (HRP)-conjugated goat anti-rabbit IgG secondary antibody (1:4000; sc-2054; Santa Cruz Biotechnology). The membranes were developed with SuperSignal West Pico Chemiluminescence Substrate (Thermo Scientific, USA) followed by visualizing with Gel Logic Imaging System (Kodak, USA). An equal loading in the lanes was evaluated by probing with a rabbit monoclonal anti-GAPDH antibody (1:4000; 14C10; Cell Signaling Technology) for total protein fraction, and rabbit polyclonal anti-RbAp46 antibody (1:1000; R3779; Sigma) for nuclear extract. Densitometric analysis was performed to calculate relative optical density of examined proteins by Kodak Imaging System software.

RNA extraction and reverse transcription. Total RNA was extracted using GeneMATRIX Universal RNA Purification Kit (EURx) after 48h post transfection, followed by DNase treatment (Sigma, USA). The reverse transcription was carried out using M-MLV Reverse Transcriptase (Promega), according to the manufacturer’s instructions. Experiments were performed twice including subsequently described procedures, maintaining the same conditions.

Figure 1. Scheme of DNA constructs generation.
Restriction enzymes used are shown. (A) GFP-SNAIL construct was linked with EcoRI enzyme by sticky ends ligation. (B) GFP-P2A-SNAIL was connected with self-cleavage P2A motif with Xhol and EcoRI enzymes, respectively. (C) Final constructs were reamplified with attB flanking primers to make site specific recombination with expression vector possible.
Fluorescent microscopy and nuclei staining. Cell nuclei were staining with Hoechst 33342 (1:1000) for 1 h before the end of experiment. Fluorescent molecules intracellular localization were evaluated using Olympus IX70 microscope after 48h post transfection.

RESULTS

Transgene expression levels

No significant difference in GFP and SNAI1 gene expression was observed between GFP-SNAI1 and GFP-P2A-SNAI1 on mRNA level (Fig. 2A, B). These results confirmed carrying out the transfection procedure properly and suggested that P2A motif does not affect transcription process. As a negative control we provided GFP carrying vector and HEK293T WT cells, which both indicated very low mRNA level of SNAI1 and no mRNA expression of GFP. Transfection of HeLa cells with SNAI1 constructs resulted in highest mRNA level of vimentinin cells transfected with GFP-P2A-SNAI1 (Fig. 6). Transcription of all used expression vectors were driven by CMV promoter.

![Figure 2. Comparison of GFP and SNAI1 relative expression levels assessed by Real-Time PCR in transfected cells. No significant difference observed in both GFP and SNAI1 expression 48 h after transfection of HEK293T cells confirmed that transgene introduction procedure was carried out properly. Data represent the mean of 2 independent experiments +S.E.](image)

![Figure 3. Western Blot analysis of total protein levels in transfected cells. Total protein fraction isolated 48 h post transfection was separated (40 µg per lane) on SDS-PAGE electrophoresis gel and detected with anti-GFP (A) or anti-SNAIL primary antibody (B), respectively. Investigated protein quantity was normalized to the expression of GAPDH and analyzed by densitometry. Data represent relative optical density (ROD) of 2 independent experiments +S.E. Representative immunoblots are shown under the graphs.](image)
Protein expression levels

Introduced constructs resulted in one fusion or two single protein expressions in HEK293T cells (Fig. 3), although GFP-SNAIL was not detectable with anti-GFP antibody in total protein fraction (Fig. 3A). The green fluorescent protein expression in GFP-P2A-SNAIL transfected cells was slightly lower than in control GFP. Interestingly, cleaved GFP protein from GFP-P2A-SNAIL showed higher molecular weight than control GFP, due to the P2A peptide residue occurrence. Both HEK293T WT and GFP had no SNAIL protein expression, because of very low mRNA level. Predictably, GFP-SNAIL appeared as a fusion protein with molecular weight around 56 kDa, while GFP-P2A-SNAIL delivered a native form of SNAIL with mass of 29 kDa (Fig. 3B). In addition, total SNAIL relative quantity for P2A linked construct was more than 3 times greater than for fusion GFP-SNAIL. Analysis of nuclear extracts showed similar to previously described results (Fig. 4), although in contrast to total protein fraction green fluorescent protein expression was detected in all cases excluding HEK293T WT (Fig. 4A). Moreover, despite the lack of NLS signal, GFP protein was accumulated also in nucleus for both GFP and GFP-P2A-SNAIL constructs. SNAIL protein nuclear levels exhibited similar trend between GFP-SNAIL and GFP-P2A-SNAIL detected with both anti-GFP and anti-SNAIL antibodies (Fig. 4A, B).

Fluorescent molecules visualization

Fluorescent microscopy evaluation confirmed molecular results. Control GFP evenly spread in whole intracellular space due to protein presence in both cytoplasm and nucleus (Fig. 5). Similarly, GFP-P2A-SNAIL construct allowed for homogenous fluorescent protein accumulation inside the cell. Fluorescence intensity of GFP and GFP-P2A-SNAIL was very similar, according to total fluorescent protein quantity. These data proved high P2A peptide-mediated cleavage efficiency. Unlike the previous, GFP-SNAIL fusion protein nuclear localization was clearly visible, most frequently in circular shape. In this case fluorescence intensity was decreased in comparison to control GFP or GFP-P2A-SNAIL.

DISCUSSION

Gene transfer procedures are well optimized to maintain protein expression in broad range of cell types. We are able to manipulate transcription and translation efficiency of introduced genes. However, the need of simultaneous expression of more than one protein have been met with incomplete success. This problem concerns especially transcription factors, whose intracellular function is limited not only by nuclear transport effectiveness, but also a proper conformation of produced protein. Thus the goal of TFs successful introduction is to maximize their nuclear accumulation to enable direct interaction between it and target DNA sequence. In the other hand, we are interested in simultaneously introduction of fluorescent marker, to visualize easily the results. This provides proper distinction between transduced and untransduced cells or even allows following cell migration pathways in vivo.

Our data suggest that P2A usage as a fusion linker does not affect transcription efficiency. Obtained results shown no difference between GFP-SNAIL and GFP-P2A-SNAIL mRNA expression levels. At this point, we expected to gain similar protein expression levels. De-
Fusion genes construction determines efficient expression of introduced transcription factors

spite this, SNAIL protein amounts were different for GFP-SNAI1 and GFP-P2A-SNAI1 in total protein fractions. Importantly, GFP-SNAI1 was detected as 56 kDa fusion protein, while GFP-P2A-SNAI1 construct allowed to obtain native form of SNAIL with 29 kDa molecular mass, with just one extra proline residue at N-terminal of protein. This enabled functional domains (such as NLS) to be exposed through correct conformation of transcription factor. Neither in total fraction, nor in nucleus SNAIL protein delivered from P2A-linker construct was detected as a fusion protein. That proved high self-cleavage ability of P2A motif. Kim et al., showed that P2A cleavage efficiency in HEK293T was around 80–95% (Kim et al., 2011). Our experiment conditions did not allow to visualize uncleaved protein, most likely due to insufficient protein amount subjected to Western Blot, or lack of detection method sensitivity. Moreover, since anti-GFP is a subject to interference with multi-dimensional protein structure, we suspect limited affinity of this antibody to bind GFP-SNAI. Therefore presence of this protein in total protein fraction was not detected by anti-GFP primary antibody. Differences between detection in total and nuclear fractions may also arise from various concentration of examined protein in these two fractions.

Apar from that, nuclear quantity of SNAIL was significantly higher in GFP-P2A-SNAI1 than in GFP-SNAI1. Considering the fact that mRNA levels were similar, this leads to conclusion that protein fusion such as GFP-SNAI affects nuclear transport effectiveness of transcription factors. Molecular weight of GFP-fused

![Figure 5. Intracellular localization of GFP assessed by fluorescent microscopy in transfected cells.](image)

Nuclei were staining with Hoechst for 1 h before the end of experiment. After 48 h post transfection culture media were replaced with PBS and fluorescent molecules were visualized. Panel (A) demonstrate bright field, (B) nuclei staining, (C) green fluorescent protein of HEK293T transfected cells. The scale bar represents 50 µm.

![Figure 6. Real-Time PCR quantification of vimentin in transfected cells.](image)

GFP-P2A-SNAI1 results in the highest vimentin gene expression increase after 48 h post transfection. Data represent the changes in vimentin expression normalized to HeLa WT cells.
protein is almost 2 times greater than SNAIL native form, what makes an active transport of SNAIL transcription factor through nuclear membrane much more difficult. The second probable mechanism of this limitation is partial coverage of SNAIL’s nuclear localization signals, while N-terminal of this transcription factor is fused with C-terminal of GFP, so that specific importins are not able to bind properly within target domains. Therefore, GFP-SNAIL intracellular amount is partially arrested in cytoplasm, where SNAIL is quickly metabolized by ubiquitination and direct degradation by proteasome (Yook et al., 2005). Interestingly, despite the lack of nuclear localization signal (NLS), control GFP appeared to be accumulated both in cytoplasm and nucleus. This situation is caused by passive diffusion of relatively small green fluorescent protein through nuclear membrane pores. It is possible to use GFP homomultimers to prevent unspecific diffusion by multiplying fluorescent protein molecular weight (Seibel et al., 2007). In addition, fluorescent microscopy analysis confirmed previous results. In case of GFP-P2A-SNAILII released fluorescent marker spread homogenously in subcellular space, unlike fusion GFP-SNAIL which is clearly accumulated inside nucleus.

Since HEK293T is easily transfectable host-cell line, it has been used to investigate intracellular faith of the introduced constructs. However, to confirm impact of SNAIL on the regulation of EMT associated genes expression, we transfected HeLa, a human cervical carcinoma cell line. Vimentin, a major structural component of mesenchymal cells, has been chosen as a target gene for SNAIL action in cervical carcinoma cells. We observed increased vimentin mRNA levels for both GFP-SNAIL and GFP-P2A-SNAIL constructs, with highest level for the latter. Our data suggests, that higher vimentin mRNA level for GFP-P2A-SNAIL relates to differences in SNAIL nuclear localization and confirms advantage of P2A motif usage.

By comparison of two strategies for introduction a pair of genes, we clearly showed the advantage of P2A motif in nuclear protein expression level. This seems to be crucial for proper functionality of delivered transcription factors like SNAIL. All generated expression vectors enable direct production of lentiviral particles to maintain stable expression of introduced genes in cells of interest. By providing higher expression in the nucleus we ensure that SNAIL will be able to successfully bind specific DNA domains and affect downstream signal transduction. Simultaneous expression of fluorescent marker allows efficient and quick distinction between transduced and untransduced cells in vitro by applying Fluorescent Activated Cell Sorting method (FACS). Moreover, following in vitro cell migration, becomes also easier due to possibility of intravital videomicroscopy (IVVM) usage.

REFERENCES

Arbab AS, Yocum GT, Wilson LB, et al., (2004) Comparison of transfection agents in forming complexes with ferumoxides, cell labeling efficiency, and cellular viability. _J Lab Clin Med_ 144: 24–32.

Batlle E, Sancho E, Franci C, et al., (2000) The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells. _Nat Cell Biol_ 2: 84–89.

Chudakov DM, Matz MV, Lukyanov S, Lukyanov KA (2010) Fluorescent proteins and their applications in imaging living cells and tissues. _Physiol Rev_ 90: 1103–1163.

Gao SY, Jack MM, O’Neill C (2012) Towards optimising the production of and expression from polycistronic vectors in embryonic stem cells. _PLoS One_ 7: e48668.

Guo Q-S, Zhu M-Y, Wang L, et al., (2012) Combined transfection of the three transcriptional factors, PDX-1, NeuroD1, and MafA, causes differentiation of bone marrow mesenchymal stem cells into insulin-producing cells. _Exp Diabetes Res_ 2012: 672013.

Kim JH, Lee S-R, Li L-H, et al., (2011) High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. _PLoS One_ 6: e18556.

Lee JM, Dedhar S, Kalluri R, Thompson EW (2006) The epithelial-mesenchymal transition: new insights in signaling, development, and disease. _J Cell Biol_ 172: 973–981.

Mokejs M (2006) IRESite: the database of experimentally verified IRES structures (www.iresite.org). _Nucleic Acids Res_ 34: D125–130.

Radcliffe PA, Mitrophanous KA (2004) Multiple gene products from a single vector: “self-cleaving” 2A peptides. _Gene Ther_ 11: 1673–1674.

Seibel NM, Eljouni J, Nalaskowski MM, Hampe W (2007) Nuclear localization of enhanced green fluorescent protein homomultimers. _Anal Biochem_ 368: 95–99.

Szmyczak AL, Workman CJ, Wang Y, et al., (2004) Correction of multi-gene deficiency in vivo using a single “self-cleaving” 2A peptide-based retroviral vector. _Nat Biotechnol_ 22: 589–594.

Yamasaki H, Sekimoto T, Ohkubo T, et al., (2005) Zine finger domain of Snail functions as a nuclear localization signal for importin beta-mediated nuclear import pathway. _Genes Cells_ 10: 455–464.

Yeh JE, Toniaso PA, Frank DA (2013) Targeting transcription factors: promising new strategies for cancer therapy. _Cancer Biol Ther_ 25: 652–658.

Yook H, Li X-Y, Ota I, et al., (2005) Wnt-dependent regulation of the E-cadherin repressor snail. _J Biol Chem_ 280: 11740–11748.