DEVELOPMENT OF AUTONOMOUSLY REPLICATING VIRAL RNA TO EXPRESS THE RECOMBINANT HUMAN GRANULOCYTE COLONY-STIMULATING FACTOR

Makhabbat Kaliyeva[1], Alexander Shustov[2]

Abstract: Recombinant human granulocyte colony-stimulating factor (G-CSF) produced in cultured mammalian cells undergoes proper posttranslational modifications and, thereby, possesses better pharmacological properties in comparison with the homologous protein expressed in bacteria. Biopharmaceuticals derived from cell culture tend to be expensive because of lower yields compared to bacterially expressed competitors and numerous issues with the scalability of production. Particular limitation of scalability pertains to delivery of expression vectors to the cell culture. Natural and efficient way to deliver foreign DNA or RNA to cell is a viral infection. We intended to develop the viral genome capable of G-CSF expression.

Objectives: To develop autonomously replicating viral RNA capable of heterologous expression of the G-CSF in cultured mammalian cells.

Methods: Genetic engineering, cell culture, and virology.

Results: Viable genome of the Venezuelan equine encephalomyelitis virus (VEE) was constructed. The G-CSF gene was synthesized de novo. Gene cassette GFP-2A-G-CSF was constructed for simultaneous expression of the fluorescent marker of viral replication (GFP) and the product of interest. Recombinant viral genome VEE/GFP-2A-G-CSF was assembled; its viability was confirmed upon transfection to the BHK-21 cells.

Conclusion: Genetic engineering of the cDNA copies of the genomes of RNA viruses provides excessive opportunities to develop mammalian expression systems. We achieved production of recombinant proteins in cultured mammalian cells transfected with the VEE viral RNA carrying the inserted genes of marker protein and G-CSF.

UDC Classification: 604.6, **DOI:** http://dx.doi.org/10.12955/cbup.v3.629

Keywords: RNA virus, autonomously replicating RNA, replicon, Venezuelan equine, encephalomyelitis virus, granulocyte colony-stimulating factor, eukaryotic expression

Introduction

In the production of biopharmaceuticals, recombinant pharmacologically active proteins became an important segment of high-tech economy with great potential for further expansion. Industrial pharmaceutical biotechnology offers drugs required in medicine and research, which either cannot be obtained in other way besides heterologous expression, or their isolation from natural sources incurs high economic costs. For the large-scale production of recombinant proteins, the key features of the suitable expression system (and, in a broader sense, technology platform) include scalability of technology, yields of the protein product, and costs of production. Escherichia coli is the most commonly used host for the production of recombinant proteins including the industrial technologies. However, some valuable proteins cannot be produced in a biologically active form in bacteria. These include, among others, the erythropoietin and the anti-hemophilic factor VIII, which are produced for medical use in the mammalian cell cultures. For other pharmacologically active proteins, such as G-CSF, a variety of production systems are available—employing expression in bacteria, yeast, plants, and cultured animal cells. Still, expression in homologous host often provides product with better properties for the end-user. For example, one form of the G-CSF produced in industrial scales in the E.coli is present in the market with an international non-proprietary name (INN) “Filgrastim” and a large number of patented brand marks. Different form of G-CSF is produced in the in vitro cultured CHO cells, which are transformed with the eukaryotic expression plasmid. This kind of G-CSF has INN “Lenograstim,” whereas Filgrastim lacks glycosylation and bears non-natural formylmethionine

[1] Makhabbat Kaliyeva, staff researcher, National Center for Biotechnology, Astana, Kazakhstan, balmakhabbat@mail.ru

[2] Alexander Shustov, PhD, head of laboratory, National Center for Biotechnology, Astana, Kazakhstan, shustov@biocenter.kz

residue at the N-terminus. Lenograstim is fully equivalent to the 174 aa natural human G-CSF. Lenograstim has greater potency compared to Filgrastim, probably because it is more stable to proteolysis (Ataergin et al., 2008).

Expression vectors constitute crucial parts of the technological platforms for production of recombinant proteins. The vectors are DNA or RNA molecules used as vehicles to carry the genes of interest (GOI), target the GOI into the expression host, and ensure GOI translation. At present, industrial production of biopharmaceuticals in cell cultures relies almost exclusively on the use of eukaryotic plasmids (episomes) as expression vectors. On the other hand, the effective systems were developed for temporary (transient) expression and provide substantially higher yields of the expression products (Núñez et al., 2013).

Efficient expression of heterologous GOIs was shown for autonomously replicating RNA-based constructs. Such RNAs are capable of autonomous replication in the cytoplasm of eukaryotic cell. These RNAs (RNA replicons) are produced by genetic engineering of genomes of RNA-containing viruses. Among diverse families of RNA viruses with cytoplasmic replication, alphaviruses (i.e., representatives of the genus Alphavirus, family Togaviridae) have several advantages as vectors (Bernal, 2013). Firstly, the alphaviruses are capable of infecting a wide range of hosts and replicating in cells of different species origin, including cells from mammals, birds, and insects (Xiong et al., 1989). Secondly, alphaviral replication machinery produces high levels of viral RNAs and proteins. It has been demonstrated that in some vertebrate cell types, Sindbis virus (SIN) produces ~$5x10^5$ molecules of viral subgenomic RNA (23S RNA) per cell, also ~$7\text{-}10x10^8$ molecules of structural protein per cell, and ~10^3 infectious particles per cell. Thirdly, genetic engineering based on manipulations with cDNA copies of the viral genomes are well developed for a number of alphaviruses, including the Venezuelan equine encephalomyelitis virus (VEE) (Davis et al., 1991), Sindbis virus (SIN) (Rice, Levis, Strauss, & Huang, 1987), Semliki Forest virus (SFV) (Liljestrom, Lusa, Huylebroeck, & Garoff, 1991), and Ross River virus (Kuhn, Niesters, & Strauss, 1991). Available knowledge on the molecular biology of these viruses allows the usage of known mutations in the non-structural proteins, in capsid protein and in non-coding regions of viral RNA to control the speed of replication, cytotoxicity, and levels viral protein biosynthesis (Hahn, Strauss, & Strauss, 1989). Efficient systems for transient protein expression utilize modified viral genomes, which are replication-competent but incapable of forming infectious virions. The latter property is because, in these genomes (replicons), the viral genes encoding structural proteins were replaced with the heterologous GOI (Xiong et al., 1989).

In this article, we describe the VEE genome designed for transient simultaneous expression of two proteins, G-CSF and a fluorescent marker (GFP) in infected or transfected cell lines.

Materials and methods

Strains and synthetic genes

To produce viral replicon, a plasmid carrying the full-length cDNA copy of the genome of VEE strain TC-83 (Genbank Acc# L01443) was used. In the original plasmid, the VEE genome is cloned downstream of promoter for SP6 RNA polymerases. This design allows using the in vitro transcription reaction to produce viral RNA for subsequent transfection. Sequence of human G-CSF gene was extracted from Genbank (Acc# CAA27291). The G-CSF gene with a natural signal peptide for transport into endoplasmic reticulum was produced using the de novo gene synthesis. eGFP gene is from laboratory collection.

Design of viral replicon

Autonomously replicating fragment of the VEE genome was produced by deleting of viral genes encoding structural proteins (C-E3-E2-6k-E1). To make heterologous expression possible, open reading frame of the viral subgenomic (26S) RNA, which encodes structural polyprotein, was completely replaced with the gene cassette GFP-2A-SP-G-CSF. Other parts of the viral genome including untranslated regions (5'-UTR and 3'-UTR) and subgenomic promoter, from which the 26S RNA is synthesized, were not subject to any changes. The gene cassette GFP-2A-SP-G-CSF was assembled by ligation of the protein-coding sequences of genes for green fluorescent protein (eGFP), autoprotease 2A of the foot-and-mouth disease virus (FMDV) and SP-G-CSF (the latter is G-CSF with native signal peptide). The gene cassette is designed to encode a polyprotein, which during translation (co-translationally) is cleaved in two parts (GFP-2A and SP-G-CSF) because of the 2A autoprotease activity. The presence of GFP among the expression products allows for easy microscopic examination of the viability of the viral RNA upon transfection. The GFP fluorescence also indicates translation of the polyprotein from gene cassette. The signal peptide directs the preprotein SP-G-CSF into an endoplasmic reticulum, where SP is cleaved off by signal peptidase, and G-CSF passes through the secretory pathway and is released into the extracellular medium. The resulting plasmid construct and replicon were designated as VEErep/GFP-2A-SP-G-CSF.

All genetic engineering was performed using standard methods of fusion PCR and cloning using restriction sites. Sequences of all intermediate constructs and the final plasmid VEErep/GFP-2A-SP-G-CSF were confirmed by bidirectional sequencing.

In vitro RNA synthesis and transfection of cell culture

For in vitro transcription, the plasmid VEErep/GFP-2A-SP-G-CSF was linearized at a unique site MluI, which is located downstream of the oligo-A stretch, which is a 3’-terminal sequence element of the viral 3'-UTR. Replicon RNA was synthesized in a reaction catalyzed by SP6 RNA polymerase; integrity of RNA and its conformity to expected size was checked using electrophoresis in agarose gel. The transcription mixture was used to transfect cultured cells without further purification. In every transfection experiment, monolayer of BHK-21 cells with 50% confluency grown in T25 flasks was used. These cells were transfected with ~4 micrograms of replicon RNA using procedure with the Lipofectamine 2000.

Results

Replicon designed for co-expression of marker protein GFP and G-CSF

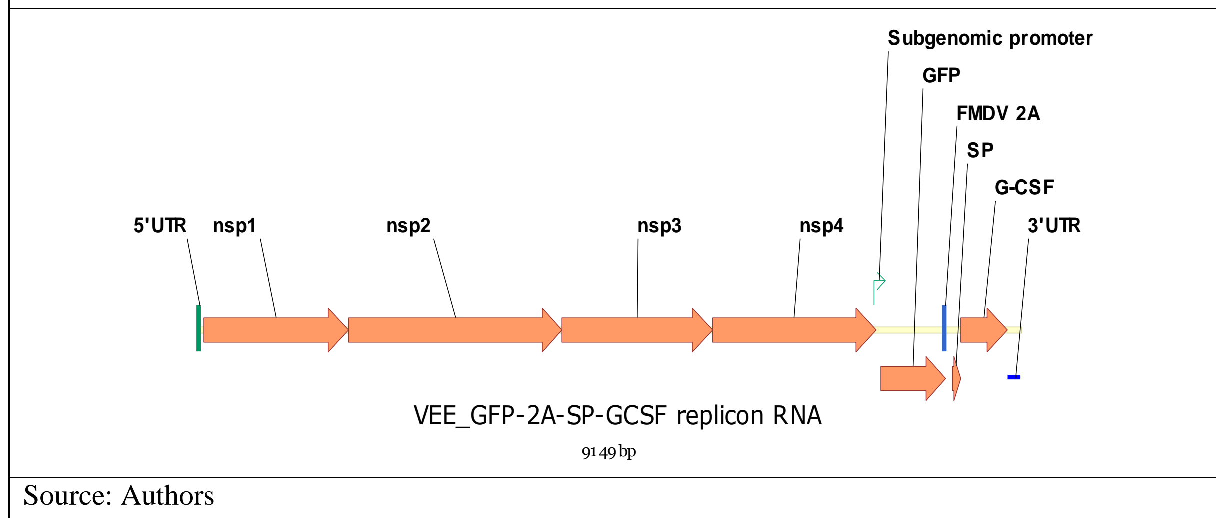
Genetic map for the replicon VEErep/GFP-2A-SP-G-CSF is shown in Figure 1; genetic map of related plasmid from which the replicon is produced is shown in Figure 2; and, amino acid sequence of the polyprotein GFP-2A-SP-G-CSF is shown in Figure 3.

In plasmid VEErep/GFP-2A-SP-G-CSF, the first nucleotide of viral genome (A) is fused to the first nucleotide (G) of transcription products synthesized by RNA polymerase SP6.

In order to conveniently control viability and monitor replication of the viral genome, we included gene encoding fluorescent marker GFP into the replicon design. For this purpose, the genetic cassette was cloned under the control of the viral 26S RNA promoter. The genetic cassette represents the open reading frames of the eGFP; autoprotease 2A and G-CSF fused end-to-end in this order. Accordingly, it encodes the polyprotein containing sequence of autoprotease 2A, which separates GFP and G-CSF. The autoprotease 2A is a 17 aa sequence (NFDLLKLAGDVESNPGP), which during translation of the polyprotein interacts with the eukaryotic ribosome in such a way that the pre-synthesized part of polyprotein (GFP-2A) is released from the ribosome, but translation continues and the remaining part

of the polyprotein (SP-G-CSF) is synthesized. Thus, two protein products are produced encoded in different parts of the same ORF. Besides fluorescent marker, different product of expression represents the product of interest—the recombinant human G-CSF.

Figure 1: Map of the autonomously replicating RNA carrying genes GFP, FMDV 2A, and G-CSF assembled into one contiguous ORF.



Source: Authors

Figure 2: Map of plasmid VEErep/GFP-2A-SP-G-CSF.

Source: Authors

Figure 3: Amino acid sequence of polyprotein GFP-2A-SP-G-CSF expressed from 26S subgenomic RNA.

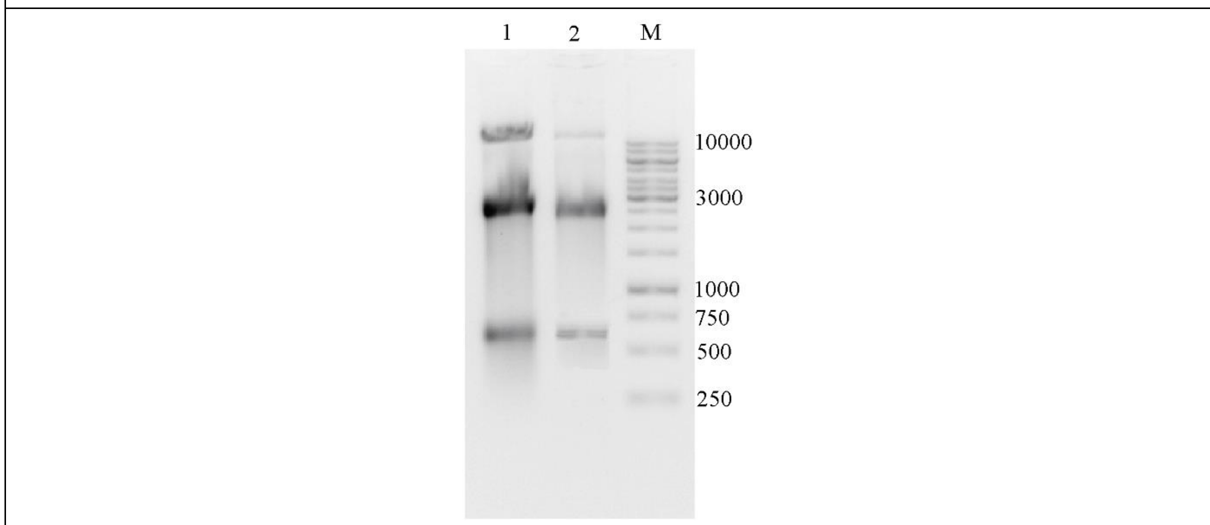
```
                                                          GFP
  1  MVSKGEELFT GVVPILVELD GDVNGHKFSV SGEGEGDATY GKLTLKFICT TGKLPVPWPT LVTTLTYGVQ CFSRYPDHMK QHDFFKSAMP EGYVQERTIF
                                                          GFP
101  FKDDGNYKTR AEVKFEGDTL VNRIELKGID FKEDGNILGH KLEYNYNSHN VYIMADKQKN GIKVNFKIRH NIEDGSVQLA DHYQQNTPIG DGPVLLPDNH
                                                                                                   G-CSF (humar
                GFP                        2A protease (FMDV)                   SP
201  YLSTQSALSK DPNEKRDHMV LLEFVTAAGI THGMDELYKL NFDLLKLAGD VESNPGPFAA AAATMAGPAT QSPMKLMALQ LLLWHSALWT VQEATPLGPA
                                                     G-CSF (human)
301  SSLPQSFLLK CLEQVRKIQG DGAALQEKLC ATYKLCHPEE LVLLGHSLGI PWAPLSSCPS QALQLAGCLS QLHSGLFLYQ GLLQALEGIS PELGPTLDTL
                                        G-CSF (human)
401  QLDVADFATT IWQQMEELGM APALQPTQGA MPAFASAFQR RAGGVLVASH LQSFLEVSYR VLRHLAQP
```

Source: Authors

Examination of viability of replicon

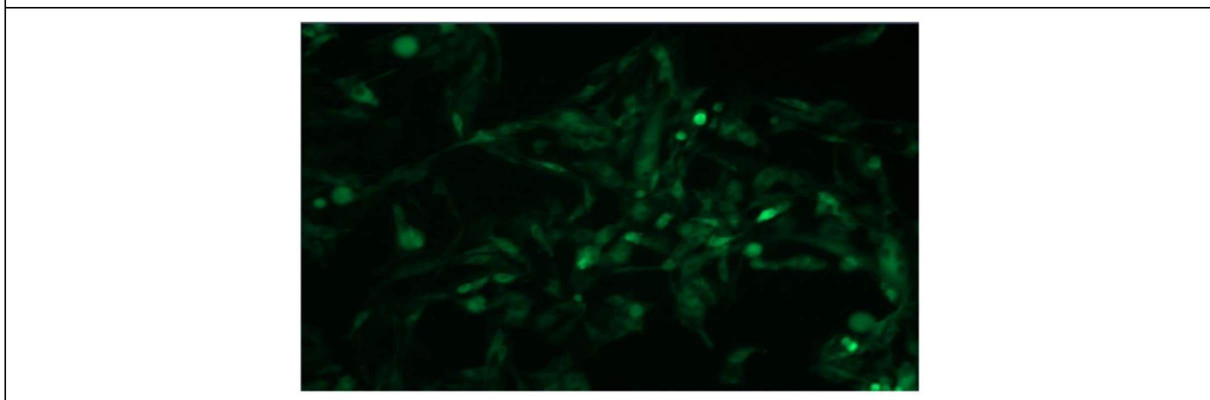
The results of in vitro RNA synthesis are presented in Figure 4.

Figure 4: Results of in vitro synthesis of replicon RNA. Lanes #1 and #2 represent two reactions. Products of both reactions contain bands ~3000 bp, which are synthesized full-length replicon RNA. The seemingly low Mw (~3000 bp) is because the single-stranded RNA adopts compact tertiary structure with mobility higher than that of dsDNA fragments with the same Mw.



Source: Authors

Figure 5: GFP fluorescence resulting from transfection of the replicon RNA into BHK-21 cells. Culture photographed at 24 hours post-transfection.



Source: Authors

Cultured cells of baby hamster kidney (BHK-21) were transfected with the replicon RNA. This experiment confirmed viability of the replicon and its ability to drive translation from the heterologous expression cassette. Majority of cells in culture demonstrated characteristic green fluorescence of GFP at 24 hours post transfection (Figure 5).

Discussion

Expression systems utilizing autonomously replicating fragments of alphaviral genomes were successfully used in laboratory experiments for the production of human transferring receptor, mouse dihydrofolate reductase, lysozyme, and beta-galactosidase. Using these systems, yields of recombinant proteins were achieved up to 25% of the total cellular protein (Liljestrom & Garoff, 1991). It was demonstrated that lengths of heterologous inserts, which remain stable during replication of viral genomes, may be up to 5 kb. Such capacity of alphaviral vectors is more than enough to develop expression constructs for production of all biopharmaceutical proteins.

Granulocyte colony-stimulating factor (G-CSF) is a major regulator of the proliferation and differentiation of neutrophilic granulocytes. Industrially produced G-CSF is widely used for treatment of immunosuppressive conditions and neutropenia (low levels of neutrophils) and for restoring hematopoiesis after chemotherapy or radiotherapy. For hospitalized patients recovering from surgery or chemotherapy, the use of G-CSF is extremely important to prevent bacterial infections (Triozzi et al., 2012; Boubeva et al., 2012). In the US, sales of G-CSF (all brands) in 2010 amounted to 605 million dollars. At present, most brands of G-CSF are produced in bacteria (Triozzi et al., 2012; Boubeva et al., 2012; Li et al., 2012). Production of G-CSF in cultured mammalian cells provides the product with better pharmacological properties. Glycosylation, which is a characteristic feature of the eukaryotically-produced G-CSF, increases the half-life of this cytokine in a systemic circulation exhibiting positive effect on pharmacokinetics. Besides, G-CSF produced in mammalian cells has correctly processed N-terminus and is less immunogenic compared to the bacterially expressed protein.

Conclusion

We created the replicon of RNA-containing virus with cytoplasmic replication, which is designed for production of recombinant proteins, among which the GFP and the commercially important biopharmaceutical product—G-CSF. Our work demonstrates a promising approach for the development of vectors for transient expression of recombinant proteins in cultured mammalian cells.

References

Ataergin, S., Arpaci, F., Turan, M., Solchaga, L., Cetin, T., Ozturk, M., Ozet, A., Komurcu, S., & Ozturk, B. (2008). Reduced dose of lenograstim is as efficacious as standard dose of filgrastim for peripheral blood stem cell mobilization and transplantation: a randomized study in patients undergoing autologous peripheral stem cell transplantation. *American Journal of Hematology, 83*(8), 644-648.

Bernal, J. A. (2013). RNA-based tools for nuclear reprogramming and lineage-conversion: towards clinical applications. *Journal of Cardiovascular Translational Research, 6*(6), 956-968.

Boubeva, R., Reichert, C., Handrick, R., Müller, C., Hannemann, J., & Borcharda, G. (2012). New expression method and characterization of recombinant human granulocyte colony-stimulating factor in a stable protein formulation. *CHIMIA International Journal for Chemistry, 66*(5), 281-285.

Davis, N. L., Powell, N., Greenwald, G. F., Willis, L. V., Johnson, B. J., Smith, J. F., &Johnston, R. E. (1991). Attenuating mutations in the E2 glycoprotein gene of Venezuelan equine encephalitis virus: construction of single and multiple mutants in a full-length cDNA clone. *Virology*, *183*(1), 20-31.

Hahn, Y. S., Strauss, E. G., &Strauss, J. H. (1989). Mapping of RNA - temperature-sensitive mutants of Sindbis virus: assignment of complementation groups A, B, and G to nonstructural proteins. *Journal of Virology, 63*(7), 3142-3150.

Kuhn, R. J., Niesters, H. G., Hong, Z., &Strauss, J. H. (1991). Infectious RNA transcripts from Ross River virus cDNA clones and the construction and characterization of defined chimeras with Sindbis virus. *Virology, 182*(2), 430-441.

Li, M., Fan, H., Liu, J., Wang, M., Wang, L., &Wang, C. (2012). High pH solubilization and chromatography-based renaturation and purification of recombinant human granulocyte colony-stimulating factor from inclusion bodies. *Applied Biochemistry and Biotechnology, 166*(5), 1264-1274.

Liljestrom, P., & Garoff, H. (1991). A new generation of animal cell expression vectors based on the Semliki Forest virus replicon. *Biotechnology (NY), 9*(12), 1356-1361.

Liljestrom, P., Lusa, S., Huylebroeck, D., & Garoff, H. (1991). In vitro mutagenesis of a full-length cDNA clone of Semliki Forest virus: the small 6,000-molecular-weight membrane protein modulates virus release. *Journal of Virology, 65*(8), 4107-4113.

Núñez, E. G. F., Jorge, S. A. C., Astray, R. M., de Rezende, A. G., da Costa, B. L. V., Monteiro, D. C. V., Pereira, C. A., & Tonso, A. (2013). Semliki forest virus as a vector: pros and cons for its use in biopharmaceuticals production. *Brazilian Archives of Biology and Technology, 56*(5), 859-866.

Rice, C. M., Levis, R., Strauss, J. H., &Huang, H. V. (1987). Production of infectious RNA transcripts from Sindbis virus cDNA clones: mapping of lethal mutations, rescue of a temperature-sensitive marker, and in vitro mutagenesis to generate defined mutants. *Journal of Virology,61*(12), 3809-3819.

Triozzi, P. L., Achberger, S., Aldrich, W., Elson, P., Garcia, J., & Dreicer, R. (2012). Differential immunologic and microRNA effects of 2 dosing regimens of recombinant human granulocyte/macrophage colony-stimulating factor. *Journal of Immunotherapy,35*(7), 587-594.

Xiong, C., Levis, R., Shen, P., Schlesinger, S., Rice, C. M., &Huang, H. V. (1989). Sindbis virus: an efficient, broad host range vector for gene expression in animal cells. *Science, 243*(4895), 1188-1191.