2A - The “go-to” Technology for Transgene Co-expression

Ekaterina Minskaia* and Garry A. Luke

1 Institute of Molecular Medicine, Av. Prof Egas Moniz, 1649-028, Lisbon, Portugal
2 Biomedical Sciences Research Complex, North Haugh, University of St Andrews, St. Andrews, Fife, Scotland

Abstract

In order to co-express multiple genes for biotechnological and biomedical applications, several approaches have been used with varying degrees of success. Currently, internal ribosome entry site (IRES) elements and “self-cleaving” 2A peptides are the most widely used. The length of the IRES can be prohibitive and IRES-dependent translation of the second open reading frame is often significantly reduced. 2A peptides have gained in popularity due to their small size and ability to consistently produce discrete proteins at an equal level. Here, we promote the use of these sequences as the “go-to” technology for co-expression of multiple proteins.

Keywords: Protein co-expression; 2A; Biotechnology; Biomedicine

Commentary

Many biotechnological and biomedical applications rely on the effective co-expression of multiple proteins. So far, multiple genes have been expressed via: (i) monocistronic vectors e.g. viral co-infection or co-transfection of plasmids expressing one protein each; fusion proteins; fusion proteins incorporating proteinase cleavage sites and (ii) polycistronic vectors where multiple genes are assembled either under the control of multiple promoters or a single promoter. IRES sequences have been used as a method to separate two coding sequences under the control of a single promoter. Despite their widespread use, they are relatively large (~600 base pairs) and expression of the downstream gene is significantly less efficient than the upstream gene [1]. A different approach using the 2A oligopeptide sequence allows multiple discrete proteins to be synthesized from a single strand of RNA, which also functions as a messenger RNA (mRNA). Not only is the 2A sequence smaller (54-174bp) than IRES elements, co-expression of proteins to be synthesized from a single strand of RNA, which also functions as a messenger RNA (mRNA). Not only is the 2A sequence smaller (54-174bp) than IRES elements, co-expression of proteins linked via 2A is independent of the cell type (cleavage activity is only linked via 2A is independent of the cell type (cleavage activity is only smaller (54-174bp) than IRES elements, co-expression of proteins linked via 2A is independent of the cell type (cleavage activity is only

IRES and 2A sequences have been used successfully in an impressive array of studies: at least 200 in the case of IRES elements and almost 900 in the case of 2A peptides making 2A the “go-to” option for protein co-expression [6-9]. Here we provide a short update on the history of 2A and cover some applications of 2A co-expression technology.

The Foot-and-mouth disease virus (FMDV) 2A sequence (hereafter “F2A”) mediates “self-processing” by a novel translational effect variously referred to as ‘ribosome skipping’, ‘stop-go’ and ‘stop-carry on’ translation [10-12]. 2A peptide cleavage has been studied in various cell types using various recombinant polyproteins and artificial reporter polyprotein systems comprising chloramphenicol acetyltransferase (CAT), β-glucuronidase (GUS), and fluorescent proteins (FPs e.g. GFP, RFP, YFP) [13-16]. It was demonstrated that F2A plus the N-terminal proline of the downstream gene is significantly less efficient than the upstream gene [1]. A different approach using the 2A oligopeptide sequence allows multiple discrete proteins to be synthesized from a single strand of RNA, which also functions as a messenger RNA (mRNA). Not only is the 2A sequence smaller (54-174bp) than IRES elements, co-expression of proteins linked via 2A is independent of the cell type (cleavage activity is only

Finally, it should be noted that (i) 2A remains as a C-terminal extension of the upstream gene, and (ii) proline forms the N terminus of the downstream gene. The presence of N-terminal proline does not seem to affect proteins which are metabolically stable [44], but when the authentic C-terminus is required for activity or subcellular targeting of certain proteins, they should either be encoded at the C-terminus of the polyprotein or followed by cleavage sequences of the mammalian Kex2p homologue, furin (AspRRRR-), porcine teschovirus-1 (“P2A”) and Thossea asigna virus (“T2A”) (Table 1) [2,4,36,43].

*Corresponding author: Ekaterina Minskaia, Institute of Molecular Medicine, Av. Prof Egas Moniz, 1649-028, Lisbon, Portugal, Tel: +351 217999566; E-mail: minkayakat@hotmail.com

Received August 21, 2015; Accepted August 24, 2015; Published August 26, 2015


Copyright: © 2015 Minskaia E, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
**Biotechnological and Biomedical Applications**

F2A is the most widely used 2A sequence in plant biotechnology and has been used to target multiple proteins to various subcellular compartments [13,34,45,66] to improve disease resistance [47,48], drought-resistance [49] and nutritional value through metabolome engineering [10,50]. Vitamin A deficiency (VAD) is a major global health issue which affects hundreds of millions of people. This problem arises because rice, the staple food source in countries where VAD is prevalent, does not produce vitamin A or its precursor β-carotene, which have a number of vital functions in the body including growth.

Since 2000, researchers have been engineering a transgenic variety of rice referred to as “golden rice” ([51,52]). Engineering rice globulin promoter [50]. GR3 of transgenic PAC had a much more terms of carotenoid production.

F2A and ‘2A-like’ sequences have been used extensively in genetic engineering of T cells for adoptive cell therapies [39-41], human terms of carotenoid production.

Finally, we think ‘2A-like’ sequences are able to function both as a signal sequence and as a translational recoding element - this leads to partitioning of the translation products between two subcellular sites (dual protein targeting). We have identified some 2A-like sequences at the N-terminus of NLRs in the genome of the purple sea urchin Stronglylocentrotus purpuratus that were putative signal sequences. Constructs encoding wild-type [Sp2A-cherryFP-T2A-GFP] or a mutated, cleavage inactive form of Sp2A were used to transfect mammalian HeLa cells – with both constructs GFP was evenly distributed throughout the cell, wild-type Sp2A lead to cherryFP localization throughout the cell and mutated Sp2A acting as a signal lead to cherryFP localization in the exocytic pathway [37,61].

In conclusion, with the number of studies that have successfully used F2A and ‘2A-like’ sequences approaching 1000, these small peptides are proving to be the ‘go-to’ technology for co-expression of multiple proteins.

**References**


<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Source</th>
<th>2A/2A-like sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>F2A</td>
<td>Foot-and-mouth disease virus (FMDV)</td>
<td>-PKVQLNFDDLKLAGDESNPG-P-</td>
</tr>
<tr>
<td>E2A</td>
<td>Equine rhinitis A virus</td>
<td>-QCTYALLKLAGDESNPG-P-</td>
</tr>
<tr>
<td>P2A</td>
<td>Porcine tevovirus-1</td>
<td>-ATFSLKLAGDESNPG-P-</td>
</tr>
<tr>
<td>T2A</td>
<td>Thoeasa asigna virus</td>
<td>-EGRSLLTCGDVSNPG-P-</td>
</tr>
</tbody>
</table>

Table 1: 2A and ‘2A-like’ sequences used for protein co-expression. The −DeXenNPGP- motif conserved among 2A/2A-like sequences is shown in red.