Bioorganic & Medicinal Chemistry Letters 24 (2014) 3724–3727

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry Letters

journal homepage: www.elsevier.com/locate/bmcl

Identification of short untranslated regions that sufficiently enhance translation in high-quality wheat germ extract

Atsushi Ogawa *, Junichiro Tabuchi, Yasunori Doi

Proteo-Science Center, Ehime University, 3 Bunkyo-cho, Matsuyama, Ehime 790-8577, Japan

ARTICLE INFO

Article history:
Received 21 May 2014
Revised 26 June 2014
Accepted 1 July 2014
Available online 5 July 2014

Keywords:
Cell-free translation
Wheat germ extract
Untranslated region
Translation enhancer
Cap-independent translation

ABSTRACT

High-quality wheat germ extract (hqWGE) is very useful for the high-yield production of various types of protein. The most important key to high productivity is the design of mRNA templates. Although the design has been refined for straightforward and efficient translation in hqWGE, there is still room for improvement in untranslated regions (UTRs), especially the 3′ UTR length, because a long, cumbersome 3′ UTR is commonly used for translation enhancement. Here we examined some short viral 3′ cap-independent translation enhancers (3′ CITEs) to identify effective ones for efficient translation in hqWGE. We then combined the most effective 3′ CITE and a 5′ enhancer to further increase the translation efficiency. mRNA with the optimal short 3′ and 5′ UTRs, both of whose length was less than 150 nt, exhibited a productivity of 1.4 mg/mL in prolonged large-scale protein synthesis in hqWGE, which was comparable to that of control mRNA with a commonly-used long 3′ UTR (~1200 nt).

To evaluate the structure and properties of a protein and to further use that protein in various applications, some quantity thereof is required. *Escherichia coli* is generally used as a high-yield translation system to obtain a large amount of protein.[1] However, *E. coli* is not suited for the expression of eukaryotic proteins due to the rapid translation: a number of these proteins expressed in *E. coli* are expected to be misfolded and to thus aggregate as insoluble precipitates.[2] Although some types of eukaryotic cells may be available for precisely folding the proteins, the expression level is often not so high, despite the performance of several laborious experimental steps. Moreover, living cells limit the types of protein for expression to those that do not adversely affect cell growth. They also present a risk of biohazards and bioethical concerns. In contrast, cell-free translation systems circumvent these disadvantages of living cells.[3,4] In particular, wheat germ extract (WGE) is very useful for the expression of a broad range of proteins; not only plant proteins but also other eukaryotic, viral, and even prokaryotic proteins.[3,5] Although the productivity in cell-free translation systems is generally considered to be low, this problem has greatly been ameliorated in WGE by improving a method for extract preparation.[6] The high-quality WGE (hqWGE) prepared by the improved method typically gives several milligrams of protein per 1 mL of the reaction volume.[6] In addition, a versatile 5′ terminal sequence in the untranslated region (UTR) that is available as an alternative to the 5′ cap has been developed for the efficient translation of 5′ cap-free mRNA in this hqWGE.[5] The sequence enables us to simplify the preparation of mRNA, concretely, so that the costly and troublesome 5′ capping of mRNA which is generally required in eukaryotic translation can be omitted. According to some reports,[5,7] 5′ cap-free mRNA with a 5′ terminal GAA triplet and a subsequent short translational enhancer, such as the omega sequence from *Tobacco mosaic virus*, or the E01 sequence (including the 5′ terminal GAA) obtained via in vitro selection exhibits high-yield translation in hqWGE as capped mRNA does. However, as opposed to the 5′ UTR, it has not yet been determined how the sequence in the 3′ UTR affects translation efficiencies in hqWGE, though it is generally known that a relatively long and thus hard-to-handle 3′ UTR is required for high-yield translation.[5] We report herein an effective short sequence in the 3′ UTR (and the corresponding sequence in the 5′ UTR) that sufficiently enhances translation in hqWGE.

Although it has already been reported that uncapped 5′-GAA mRNA with an ~1600-nt 3′ UTR shows relatively high translation efficiency in hqWGE, which is comparable to that of capped mRNA with a typical poly(A) chain,[5] it is as yet unknown how the translation efficiency depends on the 3′ UTR length. Thus, we first constructed some uncapped mRNA templates that have the E01 sequence (vide supra)[7] and 3′ UTRs of various lengths to investigate in detail the effect of the 3′ UTR length on translation efficiencies in hqWGE (Fig. 1**A**). With regard to the 3′ UTR sequence, two completely different types of sequences derived

* Corresponding author. Tel./fax: +81 89 927 8450.
E-mail address: ogawa.atsushi.mf@ehime-u.ac.jp (A. Ogawa).

http://dx.doi.org/10.1016/j.bmcl.2014.07.004
0960-894X/

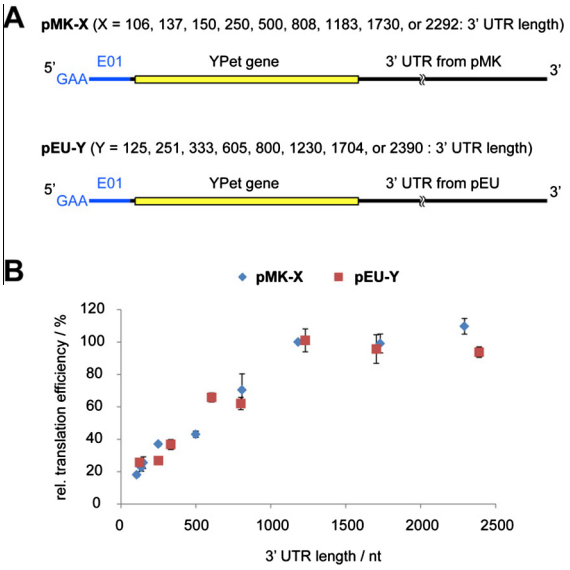

Figure 1. mRNAs with the E01 and a plasmid-derived normal sequence in the 5′ UTR and 3′ UTR, respectively. (A) Schematic diagram of YPet-encoding mRNAs, **pMK-X** and **pEU-Y**, which have 3′ UTRs of varying lengths derived from pMK and pEU, respectively. The X and Y numbers represent the 3′ UTR length. (B) The relative translation efficiencies (with standard deviations) of **pMK-X** and **pEU-Y** in hqWGE, compared to that of **pMK-1183**. They were plotted against their 3′ UTR length.

from two plasmids, pMK and pEU, were used for **pMK-X** and **pEU-Y**, respectively. We evaluated the translation efficiency of these templates, encoding yellow fluorescent protein (YPet),[8] based on the fluorescent intensity of YPet expressed with 1-h cell-free translation in batch mode in hqWGE. As a result, the efficiency increased in proportion to the 3′ UTR length up to ~1200 nt and then plateaued, in precisely the same fashion in the case of both **pMK-X** and **pEU-Y** (i.e., independently of the sequence in the 3′ UTR) (Fig. 1**B**). The reason for the proportional increase is probably that a longer 3′ UTR has much resistivity for degradation by endogenous 3′ exonucleases[5] and that it interacts more tightly with the 5′ UTR (vide infra). The plateau indicates that approx. 1200 nt of 3′ UTR is long enough to completely interact with the 5′ UTR by some mechanism (probably through some proteins) to circularize the mRNA, thereby saturating the effect of the mRNA stabilization and the ribosome recycling, as in the case of capped mRNA with a poly(A) chain.[9] In fact, this noncanonical phenomenon—circularization of mRNA that bears the 5′ terminal GAA and a long 3′ UTR instead of the 5′ cap and a poly(A) chain—has been previously observed in hqWGE by electron microscopy.[10] The above interpretation of the plateau is also consistent with the report that a 5′ GAA/long 3′ UTR (~1600 nt) pair has almost the same effect on the translation efficiency as a common 5′ cap/poly(A) chain pair does.[5]

The minimum length of the 3′ UTR for the maximum translation efficiency (i.e., ~1200 nt) is too long in comparison to the 5′ UTR length (85 nt). Such a long UTR is often an impediment in the engineering and handling of mRNA (and the template DNA). We therefore decided to identify a short 3′ UTR sequence that exerts a comparable impact on translation efficiencies. Although, judging from Figure 1**B**, there seems to be no sequence-dependency in the 3′ UTR, it should be possible to obtain a satisfactory 3′ UTR if mRNA could somehow be circularized via interactions between short 5′ and 3′ UTRs. Thus, we focused on 3′ cap-independent translation enhancers (3′ CITEs), which have been found in the 3′ UTR of various positive-strand RNA plant viral genomes.[11] A general 3′ CITE (except for those in the T-shaped structure class, vide infra) can directly interact with the 5′ UTR to circularize the mRNA

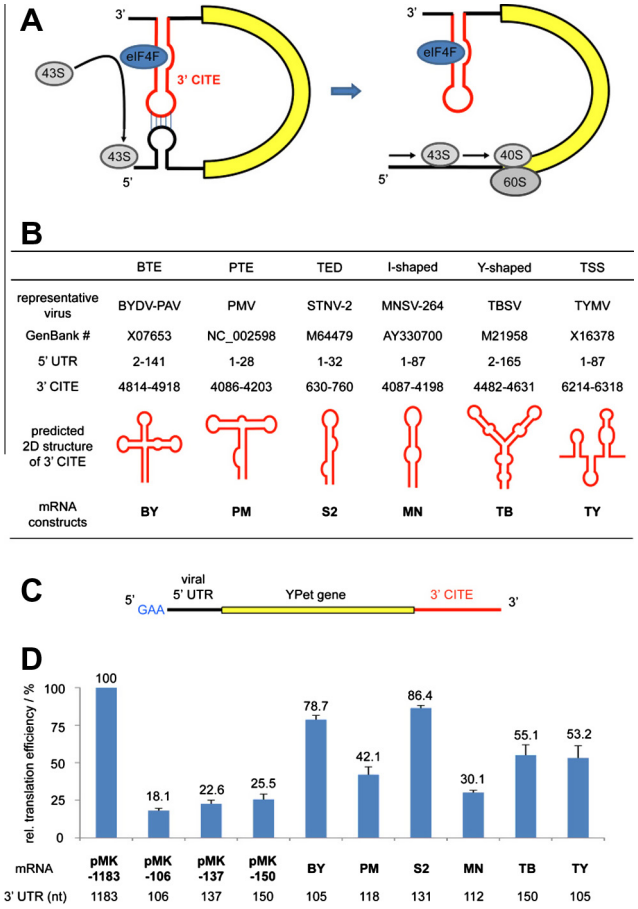

Figure 2. mRNAs with a 3′ CITE and the cognate viral 5′ UTR. (A) Proposed model of translational enhancement by typical 3′ CITEs. The mRNA is circularized by interaction between a 3′ CITE and the 3′ CITE-complementary stem-loop in the 5′ UTR without the 5′ cap and a poly(A) chain (left). The circularization and the eIF4F bound to the 3′ CITE facilitates the recruitment of the 43S ribosome to the 5′ terminus of mRNA to start translation (left to right).[11] (B) Six representative 3′ CITEs from each class. The undermost row (mRNA constructs) indicates the name of the mRNA templates with each 3′ CITE and the cognate viral 5′ UTR for translation in hqWGE (their schematic diagram is shown in C). (D) The relative translation efficiencies (with standard deviations) of the mRNAs in C in hqWGE.

without the help of both the 5′ cap and a poly(A) chain (Fig. 2**A**). In addition, this type of 3′ CITE commonly binds to the initiation factor complex eIF4F and is thus thought to recruit the 43S ribosome as the 5′ cap does. Despite the multifunctionality, the length of 3′ CITEs required for the functions is relatively short (typically 100–200 nt). In addition, 3′ CITEs are expected to show resistivity for 3′ exonucleases, even with the use of them alone in the 3′ UTR due to their rigid structure.

The 3′ CITEs reported thus far have been grouped into six classes, based on their sequence and secondary structure:[11] the *Barley yellow dwarf virus* (BYDV)-like translation element (BTE), the *Panicum mosaic virus* (PMV)-like translation element (PTE), the translational enhancer domain (TED), the I-shaped 3′ CITEs, the Y-shaped 3′ CITEs, and the T-shaped structure (TSS) (Fig. 2**B**). To widely evaluate these 3′ CITEs, we selected, from each class, one class representative that has been reported to highly enhance gene expression with a relatively short length (less than 150 nt) in plant translation systems. The selected 3′ CITEs were the BTE from BYDV,[12] the PTE from PMV,[13] the TED from *Satellite tobacco necrosis virus* (STNV)-2,[14] the I-shaped 3′ CITE from *Melon necrotic spot virus* (MNSV)-264,[15] the Y-shaped 3′ CITE from *Tomato bushy stunt virus* (TBSV),[16] and the TSS from *Turnip yellow mosaic virus* (TYMV)[17] (Fig. 2**B**). We then