Acta virologica **60**: 151 – 155, 2016 doi:10.4149/av_2016_02_151

Effect of the nucleotides surrounding the start codon on the translation of foot-and-mouth disease virus RNA

X. X. MA[1,3#], Y. P. FENG[1#], Y. X. GU[2], J. H. ZHOU,[3] , Z. R. MA[1*]

[1]College of Life Science and Engineering, Northwest University for Nationalities; Engineering & Technology Research Center for Animal Cell, Gansu; Key Laboratory of Bioengineering & Biotechnology of State Ethnic Affairs Commission; Lanzhou, 730030, P. R. China; [2]Institute of Preventive Veterinary Medicine, Zhejiang University, Hangzhou,310058, P. R. China; [3]Lanzhou Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Lanzhou, 730046, Gansu, P. R. China

Received June 15, 2015; revised November 18, 2015; accepted April 15, 2016

Summary. – As for the alternative AUGs in foot-and-mouth disease virus (FMDV), nucleotide bias of the context flanking the AUG^{2nd} could be used as a strong signal to initiate translation. To determine the role of the specific nucleotide context, dicistronic reporter constructs were engineered to contain different versions of nucleotide context linking between internal ribosome entry site (IRES) and downstream gene. The results indicate that under FMDV IRES-dependent mechanism, the nucleotide contexts flanking start codon can influence the translation initiation efficiencies. The most optimal sequences for both start codons have proved to be UUU AUG^{1st} AAC and AAG AUG^{2nd} GAA.

Keywords: foot-and-mouth disease virus; nucleotide bias; translation initiation; dicistronic reporter; internal ribosome entry site

Introduction

Foot-and-mouth disease virus (FMDV) is a member of the *Picornaviridae* family. The virus exists as seven different serotypes: O, A, C, SAT 1, SAT 2, SAT 3, and Asia l. The virus has a positive-sense RNA genome that functions like an mRNA and encodes a viral polyprotein (Knowles and Samuel, 2003). The *L* gene encoding the first component of the polyprotein is located at the 5' terminus of the whole coding sequence and possesses two potential in-frame initiation codons that generate Lab and Lb isoforms (Cao *et al*., 1995). Despite the fact that ribosomes can initiate translation of viral proteins at the first start codon, most ribosomes tend to use the second start codon as an authentic initiation codon for FMDV (Cao *et al*., 1995). López de Quinto and Martinez-Salas reported that improving the nucleotide context in this region between the two start codons did not enhance the efficiency of translation initiation at the first start codon, and the modified region had no apparent effect on the efficiency of translation initiation at the second codon either (Lopez de Quinto and Martinez-Salas, 1999). To date, it is still unclear whether nucleotide usage biases in the region flanking the initiation codons mediate alternate initiator codon selection in FMDV RNA. Previous reports have shown that this specific region may be involved in initiation codon recognition perhaps by interaction with regulating factors (Zhou *et al*., 2010, 2013). Key nucleotide positions flanking the AUG codon influence recognition of the start codon in the scanning model (Kozak, 1997). In this study, a systematic analysis of 96 FMDV genomes was conducted to identify the characteristics of the key nucleotide positions flanking the start codons in FMDV RNA. Dicistronic reporter constructs were engineered to contain the FMDV IRES element with different versions of the flanking nucleotide sequence to investigate the role of key nucleotide positions flanking the AUG start codons on translational initiation in FMDV. Based

*Corresponding author. E-mail: maxiaoxia956@163.com, mzr@xbmu.edu.cn; phone:+86-931-2938311. #These authors contributed equally to this work.

Abbreviations: CAT = chloramphenicol acetyl transferase; EGFP = enhanced green fluorescent protein; FMDV = foot-and-mouth disease virus; IRES = internal ribosome entry site

on the nucleotide usage, the optimal nucleotide context flanking each start codon proved to be UUU AUG[1st] AAC and AAG AUG[2nd] GAA.

Materials and Methods

The 96 complete RNA sequences of FMDV, which were collected from different countries at different times, were downloaded from the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/Genbank/), including 17 sequences of serotype Asia 1, 27 of serotype A, 11 of serotype C, 26 of serotype O, 5 of serotype SAT 1, 6 of serotype SAT 2, and 4 of serotype SAT 3 genomes. Their Acc. Nos are listed below: AY304994, DQ989322, DQ533483, AY687334, AY687333, FJ906802, EF614458, EF149010, DQ989321, AY593800, AY593799, AY593798, AY593797, AY593796, AY593795, EF494488, EF494487, EF494486, AY593751, AY593769, AY593789, AY593767, AY593770, AY593782, AY593783, AY593784, AY593785, AY593786, AY593790, AY593801, AY593802, AY593787, AY593788, AY593803, AY593753, AY593756, AY593757, AY593768, AY593794, AY593771, AY593758, NC_002554, FJ824812, AM409325, DQ409191, DQ409190, DQ409188, AY593809, AY593805, AY593806, AY593807, AY593808, EF552697, EF552696, EF552695, EF552694, EF552693, EF552692, EU140964, NC_004004, AJ539139, AY593819, AY593835, AY593833, AY593836, AF511039, EF175732, DQ248888, AJ320488, AJ633821, AH012984, AB079061, FJ542372, AY317098, AH012985, AY593838, AY593840, AY593841, AY593842, AY593843, AF540910, AY593847, AY593848, AY593849, AY593850, AY593851, AY593852, AY593853, AY312587, AY312589, M10975, AJ25143, NC_003992, AY593800, AY593799, AY686687.

The relative adaptiveness of a nucleotide (w_{ij}) is the ratio of usage frequency of the nucleotide to the frequency of the optimal nucleotide at the *i*-th position. To identify the 'adaptation nucleotides' flanking the AUG alternative start codon from −3 to +6 (XXX AUG XXX, X represents any nucleotide), the w_{ij} was calculated.

To determine the role of the nucleotide context flanking the start codon in alternative start codon selection in FMDV RNA, dicistronic reporter constructs were engineered to contain the FMDV IRES element with different versions of nucleotide context flanking the start codon (Fig. 1). The upstream chloramphenicol acetyl transferase (*CAT*) and downstream enhanced green fluorescent protein (*EGFP*) genes monitor 5'-cap and IRES-dependent translation, respectively. Specifically, reporter constructs contained either the IRES alone, or the IRES positioned downstream of the CAT stop codon. In the dicistronic reporter constructs, the *EGFP* gene in which the first and second codons had been deleted was fused in frame with the region downstream of the IRES element with a given nucleotide context. Plasmid pcDNA3.1 (+) was used as an expression system for *in vitro* translation. Fragments (5'-3') comprising the *Bam*HI site-*CAT* gene-IRES-UUU AUG[1st] AAC-*EGFP* gene-*Xba*I site and the *Bam*HI site-*CAT* gene-IRES-AAG AUG[2nd] GAA-*EGFP* gene-*Xba*I site (generated by Sangon Biotech (Shanghai) Co., Ltd) were inserted into plasmid pcDNA3.1(+) to construct two expression vectors, namely pCAT-IRES-AUG(1st)-EGFP and pCAT-IRES-AUG(2nd)-EGFP (Fig. 1). The IRES nucleotide sequence (GenBank Acc No. AY593835) was generated by Sangon Biotech (Shanghai) Co., Ltd. The *CAT* gene was cloned from plasmid pcDNA3.1(+)/CAT (Invitrogen). The *EGFP* gene, excluding the start codon, was cloned from plasmid pEGFP-N1.

BHK cells (Mason *et al*., 2003) were seeded into six-well plates and allowed to reach 70–90% confluence at the time of transfection with Lipofectamine, R2000 Reagent (Invitrogen) according to the protocol. BHK cells were transfected with plasmid DNA (10 μg) containing a CMV promoter using Lipofectamine, R2000 Reagent and Optimem as described previously (Lopez de Quinto and Martinez-Salas, 1999; Rose *et al*., 1991). At 12, 16, 20, 24, and 36 hr post-transfection, soluble extracts of BHK cells were prepared by lysing the cells in 120 mmol/l NaCl, 50 mmol/l Tris-HCl (pH 7.8), and 0.5% of Nonidet P-40, followed by centrifugation at 12,000 rpm for 5 min in a 1.5-ml centrifuge tube. These products were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE; 10% gels) and detected by immunoblot with a rabbit anti-CAT antibody (Sigma) and murine anti-EGFP antibody (Sigma) with horseradish peroxidase (HRP)-labeled anti-species antibodies (Sigma) and enhanced chemiluminescence (ECL; Thermo).

Results

The characteristics of nucleotide usage at the key positions flanking the start codon

By analyzing the six positions (from −3 to −1 and from +4 to +6) surrounding both start codons, a strong bias for particular nucleotides was identified. For AUG[1st], nucleotide U was preferentially used at positions −3, −2, and −1, and nucleotide A was preferentially used at positions +4 and +5 (Table 1). For AUG[2nd], nucleotides A were preferentially used at positions −3 and −2, and nucleotides G were preferentially used at position +4 (Table 2). It is interesting that the nucleotide at position −1 was always purine (G: 95/96 and A: 1/96) (Table 2).These findings suggested that these positions around AUG[2nd] play an important role in the recognition of AUG[2nd] as an authentic start codon by eukaryotic ribosomes, and that nucleotide G at position +4 neighboring AUG[2nd] might enhance the translation initiation efficiency. Based on the nucleotide usage bias for each position from −3 to +6 flanking the two start codons (Tables 1 and 2), the optimal nucleotide context flanking each start codon of FMDV RNA was determined, as follows UUU AUG[1st] AAC and AAG AUG[2nd] GAA (Fig. 1). In comparison of these two sequences, the nucleotides flanking the second start codon of FMDV RNA were more similar to the Kozak sequence.

Table 1. Frequencies and relative adaptiveness (w) of the nucleotide surrounding the first start codon

Position	A	U	C	G	W_A	W_U	W_C	W_G
-3	8	75	13	0	0.107	1.000	0.173	0
-2	11	62	21	2	0.177	1.000	0.339	0.032
-1	27	46	23	0	0.587	1.000	0.500	0
AUG	-------	-------	-------	--------	-------	--------	--------	--------
+4	88	0	0	8	1.000	0	0	0.091
+5	79	0	0	17	1.000	0	0	0.215
+6	1	31	52	12	0.019	0.596	1.000	0.231

Table 2. Frequencies and relative adaptiveness (w) of the nucleotide surrounding the second start codon

Position	A	U	C	G	W_A	W_U	W_C	W_G
-3	93	0	0	3	1.000	0	0	0.032
-2	83	0	3	10	1.000	0	0.036	0.120
-1	39	0	0	57	0.684	0	0	1.000
AUG	-------	-------	-------	--------	-------	--------	--------	--------
+4	1	0	0	95	0.011	0	0	1.000
+5	95	0	0	1	1.000	0	0	0.011
+6	75	0	0	21	1.000	0	0	0.280

The effect of the key nucleotide context flanking the start codons on gene expression

Cell extracts transfected with the two dicistronic reporter constructs (pCAT-IRES-AUG[1st]-EGFP and pCAT-IRES-AUG[2nd]-EGFP) at 12, 16, 20, 24, and 36 hr post-transfection (Fig. 2) were subjected to real-time RT-PCR. The results showed similar levels of reporter RNAs in each reaction, which indicated that the differences in reporter gene expression were due to differences in translation. The expression

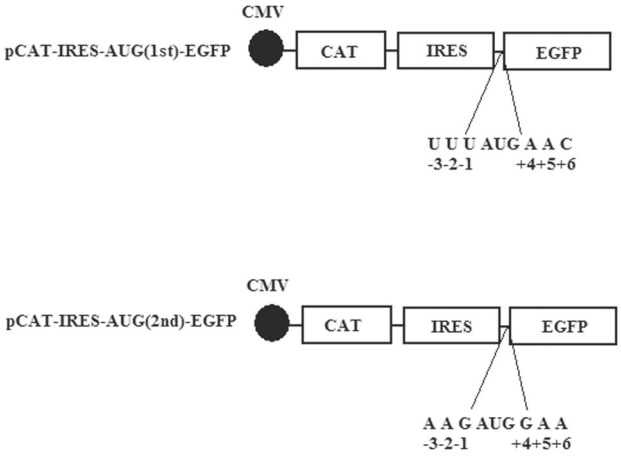

Fig. 1

Representations of the two dicistronic plasmids employed in this study
The two dicistronic reporter constructs (pCAT-IRES-AUG[1st]-EGFP and pCAT-IRES-AUG[2nd]-EGFP) with expression of CAT and EGFP under the control of a CMV promoter. Constructs were engineered to contain different versions of nucleotide context flanking the start codon. Here they are represented by optimal nucleotide context.

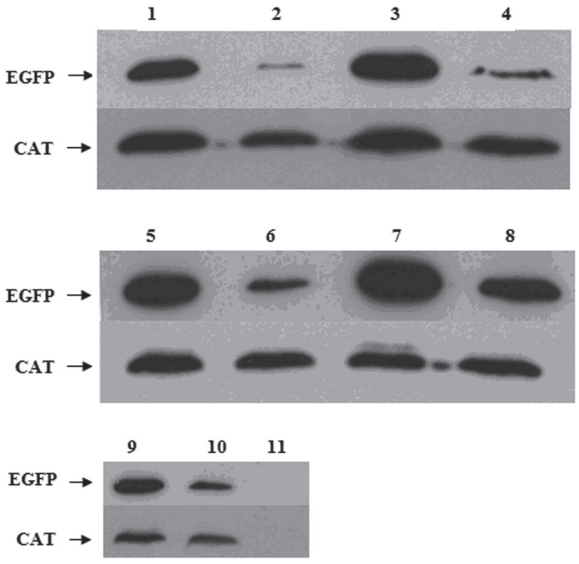

Fig. 2

The presence of optimal nucleotides flanking the start codon stimulates FMDV IRES translation
Lanes 1, 3, 5, 7, and 9 show expression of CAT and EGFP in pCAT-IRES-AUG[2nd]-EGFP-transfected BHK cells. Lanes 2, 4, 6, 8, and 10 show expression of CAT and EGFP in pCAT-IRES-AUG[1st]-EGFP-transfected BHK cells. Lane 11 represents empty plasmid serving as the negative control. Lanes 1 and 2: 12 hr post-transfection; lanes 3 and 4: 16 hr post-transfection; lanes 5 and 6: 20 hr post-transfection; lanes 7 and 8: 24 hr post-transfection; lanes 9 and 10: 36 hr post-transfection.

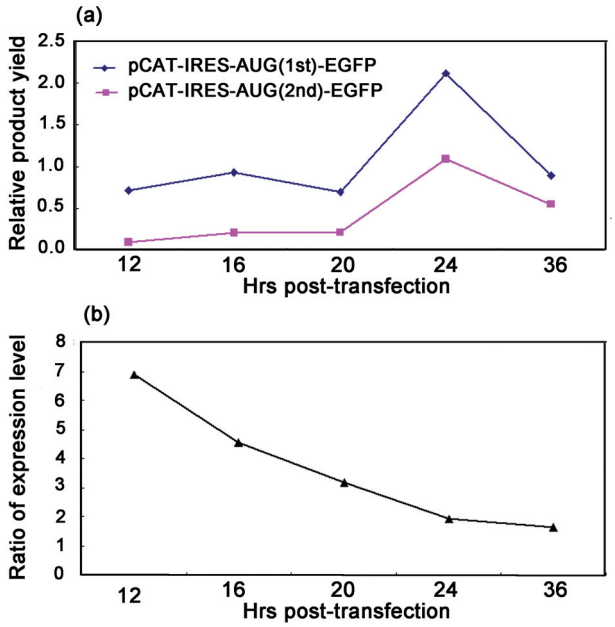

Fig. 3

Quantification of CAT and EGFP protein expression, as calculated by the EGFP/CAT ratio

IRES activity was calculated by determining the EGFP and CAT protein expression levels individually as well as the ratio of EGFP/CAT. (a) The ratio indicating the effect of the two different translation initiation regions on the expression of the *EGFP* gene. Graph showing the relative product yield. All calculations are from at least three independent experiments ± SD. (b) The ratio of the expression level of the *EGFP* gene mediated by the plasmid pCAT-IRES-AUG^{2nd}-EGFP to that of the plasmid pCAT-IRES-AUG^{1st}-EGFP at different post-translational timepoints (12, 16, 20, 24, and 36 hr).

level of the *EGFP* gene mediated by the translation initiation region (UUU AUG^{1st} AAC) was generally lower than that mediated by the translation initiation region (AAG AUG^{2nd} GAA) (Fig. 3a). It was noted that with increased time post-translation, the relative expression level of the *EGFP* gene mediated by the AAG AUG^{2nd} GAA initiation region compared with that of the UUU AUG^{1st} AAC initiation region decreased (Fig. 3b), despite the fact that the AAG AUG^{2nd} GAA initiation region mediated translation of the *EGFP* gene more effectively than the UUU AUG^{1st} AAC region. Although the exogenous gene (*EGFP* gene) does not stably exist in BHK cells, it can still be concluded that the nucleotide context flanking the start codons is crucial in mediating translation initiation in FMDV.

Discussion

The findings of this study demonstrate that the nucleotide context flanking the start codons can influence the efficiency of translation initiation of the *EGFP* gene mediated by the FMDV IRES element. A strong nucleotide bias was found in key positions (from −3 to −1 and +4 to +6) flanking both AUGs in FMDV, with the optimum sequences identified as UUU AUG^{1st} AAC and AAG AUG^{2nd} GAA. For AUG^{1st}, nucleotide U was preferentially used in positions −3, −2, and −1, with nucleotide A being preferentially used in positions +4 and +5. For AUG^{2nd}, nucleotide A was preferentially used in positions −3 and −2, and positions +4 and +5 were invariably purines (G: 95/96 and A: 1/96). A study by Grunert and Jackson similarly reported that in the case of eukaryotic translation initiation, position +5 was favored by A at an AUG or CUG start codon (Grunert and Jackson, 1994). It was also reported that recognition of AUG and alternative initiator codons is augmented by G in position +4 by direct analysis of AUG codon recognition using mRNAs with mutations in positions +4, +5, and +6 (Kozak, 1997). The purines at positions −3 and +4 are most important for maintaining the fidelity of initiation codon selection in eukaryotes (Pisarev *et al.*, 2006). Compared with different base biases in these key positions flanking the AUG^{1st}, the strong bias of nucleotide A in positions −3 and −2, G in position +4, and A in position +5 flanking the AUG^{2nd} of FMDV was more consistent with the conserved nucleotides in key positions flanking the AUG codon in eukaryotes. This finding implies that under translation selection, the base biases at specific positions flanking the AUG^{2nd} codon could act as a translation initiation signal assisting eukaryotic ribosomes in recognizing the AUG^{2nd} of FMDV. It is interesting that similar nucleotide biases at positions −2, −3, and +4, which result in high translation initiation efficiency at the start codon, were reported only in eukaryotes or other prokaryotes (Kozak, 1997; Nakagawa *et al.*, 2008; Pisarev *et al.*, 2006). This is the first report that the nucleotide bias in these key positions flanking the AUG codon also exists in viruses and that the key nucleotides are highly conserved with those reported in eukaryotes. This may be explained by the fact that the FMDV genome presents a similar AUG context to adapt to the eukaryotic translation initiation model, under translation selection. In a previous study, it was reported that substitution of nucleotides flanking the first start codon of the FMDV *Lab* gene enhanced the efficiency of translation initiation (Poyry and Jackson, 2011). Our findings elucidate that the optimal translation initiation region for the start codon of this gene could be determined by analyzing the population-evolutionary features of this nucleotide region, and that in this way the efficiency of translation initiation of the *Lab* and *Lb* genes in FMDV RNA could be influenced.

Acknowledgements. This work was supported by the Program for Changjiang Scholars and Innovative Research Team in University (IRT13091) and the National Natural Science foundation of China (No. 31302100 and No. 31160033) and the Central Universities deriving from the Northwest University for Nationalities (31920150077).

References

Cao X, Bergmann IE, Fullkrug R, Beck E (1995): Functional analysis of the two alternative translation initiation sites of foot-and-mouth disease virus. J. Virol. 69, 560–563.

Grunert S, Jackson RJ (1994): The immediate downstream codon strongly influences the efficiency of utilization of eukaryotic translation initiation codons. EMBO J. 13, 3618–3630.

Knowles NJ, Samuel AR (2003): Molecular epidemiology of foot-and-mouth disease virus. Virus Res. 91, 65–80. http://dx.doi.org/10.1016/S0168-1702(02)00260-5

Kozak M (1997): Recognition of AUG and alternative initiator codons is augmented by G in position +4 but is not generally affected by the nucleotides in positions +5 and +6. EMBO J. 16, 2482–2492. http://dx.doi.org/10.1093/emboj/16.9.2482

Lopez de Quinto S, Martinez-Salas E (1999): Involvement of the aphthovirus RNA region located between the two functional AUGs in start codon selection. Virology 255, 324–336. http://dx.doi.org/10.1006/viro.1999.9598

Mason PW, Grubman MJ, Baxt B (2003): Molecular basis of pathogenesis of FMDV. Virus Res. 91, 9–32. http://dx.doi.org/10.1016/S0168-1702(02)00257-5

Nakagawa S, Niimura Y, Gojobori T, Tanaka H, Miura K (2008): Diversity of preferred nucleotide sequences around the translation initiation codon in eukaryote genomes. Nucleic Acids Res. 36, 861–871. http://dx.doi.org/10.1093/nar/gkm1102

Pisarev AV, Kolupaeva VG, Pisareva VP, Merrick WC, Hellen CU, Pestova TV (2006): Specific functional interactions of nucleotides at key -3 and +4 positions flanking the initiation codon with components of the mammalian 48S translation initiation complex. Genes Dev. 20, 624–636. http://dx.doi.org/10.1101/gad.1397906

Poyry TA, Jackson RJ (2011): Mechanisms governing the selection of translation initiation sites on foot-and-mouth disease virus RNA. J. Virol. 85, 10178–10188. http://dx.doi.org/10.1128/JVI.05085-11

Rose JK, Buonocore L, Whitt MA (1991): A new cationic liposome reagent mediating nearly quantitative transfection of animal cells. Biotechniques 10, 520–525.

Zhou JH, Zhang J, Ding YZ, Chen HT, Ma LN, Liu YS (2010): Characteristics of codon usage bias in two regions downstream of the initiation codons of foot-and-mouth disease virus. Biosystems 101, 20–28. http://dx.doi.org/10.1016/j.biosystems.2010.04.001

Zhou JH, Zhang J, Sun DJ, Ma Q, Ma B, Pejsak Z, Chen HT, Ma LN, Ding YZ, Liu YS (2013): Potential roles of synonymous codon usage and tRNA concentration in hosts on the two initiation regions of foot-and-mouth disease virus RNA. Virus Res. 176, 298–302. http://dx.doi.org/10.1016/j.virusres.2013.06.006