



Selection of functional 2A sequences within foot-and-mouth disease virus; requirements for the NPGP motif with a distinct codon bias

Kjær, Jonas; Belsham, Graham J.

Published in:
R N A

Link to article, DOI:
[10.1261/rna.063339.117](https://doi.org/10.1261/rna.063339.117)

Publication date:
2018

Document Version
Peer reviewed version

[Link back to DTU Orbit](#)

Citation (APA):
Kjær, J., & Belsham, G. J. (2018). Selection of functional 2A sequences within foot-and-mouth disease virus; requirements for the NPGP motif with a distinct codon bias. *R N A*, 24, 12-17.
<https://doi.org/10.1261/rna.063339.117>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

1 **Selection of functional 2A sequences within foot-and-mouth disease virus; requirements**
2 **for the NPGP motif with a distinct codon bias**

3

4 Jonas Kjær¹, Graham J. Belsham¹.

5 DTU National Veterinary Institute, Technical University of Denmark, Lindholm, DK-4771

6 Kalvehave, Denmark

7

8 Short title: Codon bias for NPGP motif in FMDV 2A

9

10 Keywords: Picornavirus; Synonymous codon; Codon bias; translation

11

12 Corresponding author: Graham J. Belsham, grbe@vet.dtu.dk

13

14

15 **Abstract**

16 Foot-and-mouth disease virus (FMDV) has a positive-sense ssRNA genome including a
17 single, large, open reading frame. Splitting of the encoded polyprotein at the 2A/2B junction
18 is mediated by the 2A peptide (18 residues long) which induces a non-proteolytic, co-
19 translational, “cleavage” at its own C-terminus. A conserved feature among variants of 2A is
20 the C-terminal motif N¹⁶P¹⁷G¹⁸/P¹⁹ where P¹⁹ is the first residue of 2B. It has been shown
21 previously that certain amino acid substitutions can be tolerated at residues E¹⁴, S¹⁵ and N¹⁶
22 within the 2A sequence of infectious FMDVs but no variants at residues P¹⁷, G¹⁸ or P¹⁹ have
23 been identified. In this study, using highly degenerate primers, we analysed if any other
24 residues can be present at each position of the NPG/P motif within infectious FMDV. No
25 alternative forms of this motif were found to be encoded by rescued FMDVs after 2, 3 or 4
26 passages. However, surprisingly, a clear codon preference for the wt nucleotide sequence
27 encoding the NPGP motif within these viruses was observed. Indeed, the codons selected to
28 code for P¹⁷ and P¹⁹ within this motif were distinct; thus the synonymous codons are not
29 equivalent.

30

31

32

33 **Introduction**

34 Foot-and-mouth disease virus (FMDV) is the prototypic member of the genus *Aphthovirus*
35 within the family *Picornaviridae*. This virus is the causative agent of the highly contagious
36 and economically important disease of cloven-hoofed animals, foot-and-mouth disease. The
37 positive-sense ssRNA genome of around 8400 nt includes a single, large, open reading frame
38 (ORF), ca. 7000 nt, encoding a polyprotein (Belsham 2005). The full-length viral polyprotein
39 is never observed since it is rapidly processed during and after synthesis mainly by the virus-
40 encoded proteases (primarily 3C^{pro}) to produce 15 distinct mature proteins plus multiple
41 precursors (reviewed in Martinez-Salas and Belsham 2017). Interestingly, FMDV, like many
42 (but by no means all) other picornavirus (e.g. cardioviruses, erboviruses, teschoviruses etc.)
43 employs a co-translational, protease-independent mechanism for the “cleavage” of the
44 polyprotein at the 2A/2B junction (the boundary between the capsid proteins and the non-
45 structural proteins) (Donnelly et al. 2001a). This mechanism has been referred-to as
46 “ribosomal skipping” or, alternatively, “stop-carry on” or “StopGo” (Atkins et al. 2007;
47 Doronina et al. 2008; Donnelly et al. 2001a; Tulloch et al. 2017). The 2A peptide lacks
48 characteristic protease motifs and only mediates “cleavage” during translation. It has been
49 demonstrated that the 2A sequence is able to mediate “cleavage” in all eukaryotic translation
50 systems tested whereas a number of artificial polyproteins containing this sequence have been
51 examined in prokaryotic systems and no detectable cleavage products were observed
52 (Donnelly et al., 1997).

53 The 2A peptide contains a highly conserved D¹²(V/I)E(S/T)NPG_{2A}↓P¹⁹_{2B} motif at its C-
54 terminus which is critical for its function (Ryan and Drew 1994; Donnelly et al. 1997). This
55 motif, together with upstream amino acids, is believed to interact with the ribosomal exit
56 tunnel. This prevents the formation of a peptide bond between the C-terminal glycine (G¹⁸) of
57 2A and the N-terminal proline of 2B, referred to here, as P¹⁹ since it is an important part of

58 the cleavage mechanism (see also Donnelly et al. 2001a; Ryan et al. 1999). However,
59 remarkably, protein synthesis continues without the requirement for a re-initiation event.

60 Investigations into the activity of the 2A sequence have mainly been performed using *in vitro*
61 experiments. Typically, these have either used mRNAs with single ORFs encoding artificial
62 polyproteins comprising two reporter proteins linked via the 2A peptide (Ryan et al. 1991;
63 Ryan and Drew 1994; Donnelly et al. 2001b) or by expressing cDNAs encoding a truncated
64 viral polyprotein including the StopGo coding region (Palmenberg et al. 1992). Alterations to
65 the conserved D¹²(V/I)E(S/T)NPG_{2A}[↓]P¹⁹_{2B} motif reduced or abrogated the StopGo function
66 (Donnelly et al. 2001b; Sharma et al. 2012), thereby showing that these amino acids are
67 important for the correct StopGo “cleavage”. Furthermore, Hahn and Palmenberg (1996)
68 demonstrated that alterations to this motif also influenced the viability of
69 encephalomyocarditis virus (EMCV, a cardiovirus) as they resulted in lethal phenotypes.
70 Subsequently, Loughran et al. (2013) reported a similar observation for FMDV, as
71 modification of the S¹⁵NPG¹⁸_{2A}[↓]P¹⁹_{2B} sequence to S¹⁵NPL¹⁸_{2A}[↓]V¹⁹_{2B} or S¹⁵NPA¹⁸_{2A}[↓]P¹⁹_{2B} also
72 gave rise to a lethal phenotype.

73 However, recently, certain amino acid substitutions (e.g. 2A S¹⁵ to F/I and 2A N¹⁶ to H) that
74 have been shown to severely (60-70%) impair “cleavage” at the 2A/2B junction, using *in*
75 *vitro* assays (Donnelly et al. 2001b), have been found to be tolerated within infectious
76 FMDVs (Kjær J and Belsham GJ, submitted). In contrast, other substitutions (e.g. P¹⁹ to A
77 and P¹⁹ to G) that inhibit cleavage more severely (by 89-100%) *in vitro*, were not found
78 within rescued viruses. Indeed, viruses rescued from these mutant transcripts had sequences
79 that exactly matched the wt sequence (i.e. the rescued viruses were not mutant). In these
80 studies, we also determined a critical role for the StopGo mechanism for the overall level of
81 replication/translation of FMDV RNA. FMDV replicons with a defective 2A sequence had a

82 markedly lower replication efficiency compared to the wt replicon (Kjær J and Belsham GJ,
83 submitted).

84 It is, therefore, apparent that some amino acid substitutions can be tolerated within the
85 FMDV 2A peptide whereas other changes are not compatible with viability. To identify if
86 any alternative residues can be accepted within the critical N¹⁶P¹⁷G¹⁸_{2A}↓P¹⁹_{2B} motif,
87 degenerate sequences, encoding all possible amino acid substitutions at each of these
88 positions individually, were introduced into a full-length FMDV cDNA, as used previously
89 (Gullberg et al. 2013; Kristensen et al. 2017). In principle, this should result in the production
90 of RNA transcripts encoding 2A peptides with a wide spectrum of “cleavage” activities. This
91 was achieved by generating a large pool of plasmids, using site-directed mutagenesis with
92 highly degenerate oligonucleotides, to change each of the individual codons corresponding to
93 the amino acid residues within this conserved motif to NNN (where N is a mixture of all 4
94 bases). Using each pool of plasmids, RNA transcripts were prepared, *in vitro*, and introduced
95 into baby hamster kidney (BHK) cells. Infectious viruses were rescued and characterized.

96

97 **Results and Discussion**

98 The expected generation of a pool of StopGo cDNA mutants that could potentially result in
99 all possible single amino substitutions in place of the N¹⁶, P¹⁷, G¹⁸ and P¹⁹ residues (see
100 Figure 1A) was analysed by sequencing (see Figure 1B). The heterogeneity at the expected
101 positions was clear in each case (this does not prove that each of the possible codons was
102 present but indicates it is likely).

103 Full-length RNA transcripts were produced, *in vitro*, and introduced into BHK cells.
104 Infectious virus was generated and passaged in fresh cells. RNA was then extracted from the
105 virus harvests and the sequence encoding the 2A peptide was amplified by RT-PCR. The

106 pool of amplicons was introduced into the pCR-XL-TOPO vector and then the sequence of
107 the inserts in 20 individual colonies was determined for each virus harvest. It was found that
108 all of the rescued viruses analysed after passages p2, p3 and p4 encoded the wt amino acid
109 sequence at the NPGP motif in 2A. Interestingly, the complete spectrum of the possible
110 synonymous codons for each of the residues N¹⁶, P¹⁷, G¹⁸ and P¹⁹ was present in the rescued
111 viral genomes at p2 (see Table 1). These results indicated that the approach had indeed
112 generated a diverse pool of codons within the viruses. Furthermore, the very restricted range
113 of nucleotide sequences encoding 2A observed within the rescued viruses strongly suggests
114 that the specific amino acid sequence (NPGP), encoded by these nucleotide sequences, is
115 critical for FMDV viability.

116 However, it was also apparent that the utilization of the different codons for the conserved
117 amino acid residues varied. At p2, 55% of the sequences analysed had the wt codon for
118 residue N¹⁶ (AAC) while the synonymous AAT codon was present in the remaining 45% of
119 the rescued sequences. In the subsequent passages, the proportion of the AAC codon within
120 the sequences increased to 75% and 95% by p3 and p4 respectively while the incidence of the
121 AAT codon declined (Table 1). For residue P¹⁷, at p2, the codon CCT was present in 55% of
122 the colonies analysed and increased to 100% by p4. Each of the three other possible codons
123 for P¹⁷ (CCC, CCA and CCG) were also observed at p2 but each declined as the wt codon
124 became dominant. For residues G¹⁸ and P¹⁹, the wt codons (GGG and CCC respectively)
125 were in the minority (10 or 20%) at p2 and each of the synonymous codons were also present.
126 However, interestingly, by p3 the wt codons had markedly increased to 50% abundance and
127 by p4 were dominant ($\geq 90\%$ abundance). For G¹⁸, the GGA codon was the most abundant at
128 p2 but declined during further passages to be only 10% of the sequences at p4. Similarly, for
129 P¹⁹ the CCT codon was present in 50% of the sequences at p2 but declined to just 5% by p4.
130 Strikingly, by p3, the wt codon was present in 50-75% of the population at each of the 4

131 residues and by p4 the wt codon was present in 90-100% of the virus population in each case
132 (Table 1). Thus, it appears that selection occurs for the wt nucleotide sequence during
133 passage of the rescued viruses in cell culture.

134 The wt GGGCCC nt sequence encoding residues G¹⁸ and P¹⁹ is recognized in DNA by the
135 restriction enzyme *ApaI* (see Figure 1A). Hence, it was possible to deplete the cDNA
136 amplicons generated by RT-PCR, of the wt sequence from the rescued viruses by digesting
137 them with *ApaI* prior to the cloning step (it was anticipated that this should enhance the
138 detection of non-wt nucleotide sequences). The residual, full-length, 650bp amplicons were
139 inserted into the pCR-XL-TOPO vector, as described above, and the plasmid DNA from
140 individual colonies was sequenced. As expected, the wt codons for G¹⁸ and P¹⁹ were no
141 longer observed in the cloned fragments (Table 2) and the G¹⁸ (GGA) and P¹⁹ (CCT) codons
142 were predominant in these enriched populations. These results are consistent with those
143 obtained without the *ApaI* digestion (since the GGA and CCT codons were also present in
144 50% of the fragments at p2 without this treatment, see Table 1) but clearly the apparent
145 abundance of these non-wt codons is enhanced following the *ApaI* digestion (Table 2), as
146 anticipated. The enrichment for non-wt sequences did not result in the detection of codons for
147 alternative amino acids within the virus population. It had been anticipated that some amino
148 acid substitutions at residue N¹⁶ might be rescued since a mutant (with N¹⁶ changed to H) has
149 been shown to be viable (Kjær J and Belsham GJ, submitted) but, presumably, it was
150 outcompeted by the wt virus.

151 It is interesting to note that the G¹⁸ (GGA) and P¹⁹ (CCT) codons have previously been found
152 to be the second most abundant codons found in FMDV genomes from all seven serotypes
153 (see Gao et al. 2014). This comparison of FMDV sequences also indicated that the alternate
154 codon for N¹⁶ (AAC) is present in only a small minority of FMDV genomes and CCC is also
155 a minor population of the codons used for residue P¹⁷. The results presented in Table 1

156 clearly indicate that infectious FMDVs with these synonymous changes can be obtained but
157 these viruses do not appear to be stably maintained in cell culture and are apparently selected
158 against.

159 The evidence presented here strongly suggests that there is a distinct selection, within the
160 virus when grown in cell culture, for codon AAC for N¹⁶, CCT for P¹⁷, GGG for G¹⁸ and
161 CCC for P¹⁹; thereby indicating that synonymous codon usage for this conserved motif is
162 biased in these rescued viruses. It is particularly noteworthy that the codon preference for P¹⁷
163 and P¹⁹ is different (CCT and CCC respectively). This raises the question of why does the
164 virus select some codons over others? Various studies have demonstrated that synonymous
165 codon usage bias plays an important role in the translation of certain mRNAs (Akashi 2001;
166 Bulmer 1991; Novoa and Ribas de Pouplana 2012; Mauro and Chappell 2014). It is therefore
167 conceivable that synonymous codons may influence the cleavage efficiency through the
168 FMDV StopGo mechanism. As indicated above, a marked codon bias within the FMDV
169 genome is apparent from the alignment of diverse FMDV 2A sequences as described by Gao
170 et al. (2014). However, in the context of a synthetic reporter polyprotein, assayed within
171 CHO cells, use of the four different synonymous codons for residue G¹⁸ of the 2A peptide
172 resulted in very similar apparent “cleavage” efficiencies at the 2A/2B junction. This was
173 interpreted as showing that it is the amino acid residue rather than the nt sequence which is
174 critical for achieving cleavage (Gao et al. 2014). However, using that assay system, the
175 “cleavage” efficiency was only about 88-89% while essentially 100% cleavage occurs within
176 the native context, as in the virus. The results obtained here (see Table 1) indicate that two
177 separate selection effects may be operating. There is a clear selection for the NPGP motif at
178 the amino acid level. However, in addition, there is a distinct codon bias within the context of
179 the rescued infectious viruses and a significant selection pressure appears to exist for the wt
180 sequence. This effect is fully consistent with the codon bias observed in the analysis of

181 natural FMDV genomic sequences (Gao et al. 2014). This suggests that the FMDV RNA
182 sequence itself (rather than just the encoded amino acid sequence) affects the “cleavage”
183 process (StopGo mechanism) at the 2A/2B junction. Such an effect could be achieved
184 through a direct interaction of the RNA sequence itself or potentially through interactions
185 with the specific charged tRNAs involved in the translation process. In the case of the P¹⁷ and
186 P¹⁹ codons, it is interesting to note that the same type of prolyl tRNA (with an IGG
187 anticodon) has been reported to be used for decoding of the CCC and CCU codons in human
188 cells (no gene for a tRNA that is cognate for CCC was identified, see Mauro and Chappell,
189 2014). However, in the current database of tRNA sequences from the Lowe laboratory, it
190 appears that in humans, 1 of 23 genes for prolyl tRNAs has a GGG anticodon with 10 copies
191 having an AGG anticodon. In the mouse genome, 1 of 20 genes for the prolyl tRNAs has the
192 GGG anticodon and 8 genes have the AGG anticodon (see the gtrnadb.ucsc.edu database
193 described in Chan and Lowe (2009)). Interestingly, in cattle and pigs (major hosts for
194 FMDV) and also in the rat, there is no gene for a prolyl tRNA with a GGG anticodon. Thus,
195 it is not clear whether a single, post-transcriptionally modified prolyl tRNA recognizes these
196 two Pro codons (at least some of the time) or if different tRNAs are involved in the hamster
197 cells used here. If a single tRNA is involved in recognizing both codons (as in cattle, pigs and
198 rats), then it seems that the RNA sequence itself must be influencing the StopGo process; it
199 seems unlikely that this effect is mediated through some secondary or tertiary RNA structure,
200 as this would presumably be lost on the ribosome during the process of translation. It will
201 clearly be important to analyse the effect of the presence of the non-optimal synonymous
202 codons on “cleavage” at the 2A/2B junction in its native context.

203

204 **Materials and Methods**

205 **Construction of plasmids containing full-length mutant FMDV cDNAs**

206 Pools of StopGo cDNA mutants that potentially result in all possible single amino
207 substitutions in place of the N¹⁶, P¹⁷, G¹⁸ and P¹⁹ residues, respectively, were constructed.
208 This was achieved using a 2-step site-directed mutagenesis procedure. This is a variation of
209 the QuickChange protocol (Stratagene), using Phusion High-Fidelity DNA Polymerase
210 (Thermo Scientific) with modified versions of the plasmid pT7S3 (Ellard et al. 1999) as
211 template. The wt pT7S3 contains the full-length cDNA for the O1Kaufbeuren B64 strain of
212 FMDV. To eliminate the possibility of carrying over some residual wt template from the
213 PCR, the templates used were modified versions of the pT7S3 with the codons for N¹⁶, P¹⁷,
214 G¹⁸ or P¹⁹ changed to encode an alanine (A) residue in each case (see Figure 1A). These
215 substitutions have been reported previously to result in a complete loss of apparent cleavage
216 activity (Sharma et al. 2012; Donnelly et al. 2001b) and it has not been possible to rescue
217 infectious virus containing these substitutions (Kjær J and Belsham GJ, submitted). The first
218 round of PCRs used the forward mutagenic 2A PCR primers (Table 3), with a single reverse
219 primer 8APN206 (Table 3) plus the four different modified pT7S3 plasmids as templates and
220 generated amplicons of ca. 450 bp. These primary PCR products were then used as
221 megaprimers for a second round of PCR with the respective mutant pT7S3 plasmids as
222 templates to produce full-length plasmids. Following *DpnI* digestion, the products from each
223 reaction were introduced into *E. coli* and grown as separate pools. The plasmid pools were
224 sequenced using a BigDye Terminator v. 3.1 Cycle Sequencing Kit and a 3500 Genetic
225 Analyzer (Applied Biosystems).

226

227 **Rescue of virus from full-length cDNA plasmids**

228 Plasmid DNA isolated from each pool was linearized by digestion with *HpaI* and RNA
229 transcripts were prepared using T7 RNA polymerase (Ambion T7 MEGAscript) at 37°C for 4
230 hours. The integrity of the transcripts was assessed on agarose gels and quantified by

231 spectrophotometry (NanoDrop 1000, Thermo Scientific) after which they were introduced
232 into BHK cells by electroporation, as described previously (Nayak et al. 2005). The BHK
233 cells were maintained in Dulbecco's Modified Eagles Medium (DMEM) supplemented with
234 5% fetal calf serum, and incubated at 37°C with 5% CO₂. At 2 days post-electroporation, the
235 viruses were harvested by freezing and then amplified through three passages (p2, p3 and p4)
236 in BHK cells.

237

238 **Characterization of viruses following multiple passages**

239 After each passage, viral RNA was extracted from a sample of the virus harvest (using the
240 RNeasy Mini Kit, Qiagen) and converted to cDNA using ready-to-go you-prime first-strand
241 beads (GE Healthcare Life Sciences). FMDV cDNA, which included the whole 2A coding
242 region, was amplified in PCRs (AmpliTaq Gold DNA polymerase, Thermo Scientific) using
243 primers 8APN206 and 8APN203 (see Figure 1 and Table 3). Control reactions, without RT,
244 were used to ensure that the analysed products were derived from RNA and not from the
245 presence of carryover plasmid DNA template. The amplicons (ca. 650 bp) were visualized on
246 1% agarose gels and purified (GeneJET gel extraction kit, Thermo Scientific). These
247 amplicons should be representative of the heterogeneity present in the rescued virus
248 populations. The resulting collections of fragments were inserted into pCR-XL-TOPO
249 (Thermo Scientific) and the sequence of the cDNA fragment present in individual bacterial
250 clones (20 colonies for each of the 4 residues) was determined using the same reverse primer
251 as used for the PCR. The fragments from codon mutants G¹⁸ and P¹⁹ were also enriched for
252 the non-wt sequence populations by digestion of the cDNA with *ApaI* prior to gel purification
253 and insertion into the pCR-XL-TOPO vector as described above.

254

255 **Acknowledgements**

256 We thank Preben Normann for excellent technical assistance. The studies were supported by
257 core funds within DTU-Vet.

258

259 **References**

- 260 Akashi H. 2001. Gene expression and molecular evolution. *Curr Opin Genet Dev* **11**: 660–
261 666.
- 262 Atkins JF, Wills NM, Loughran G, Wu C-Y, Parsawar K, Ryan MD, Wang C-H, Nelson CC.
263 2007. A case for “StopGo”: reprogramming translation to augment codon meaning of
264 GGN by promoting unconventional termination (Stop) after addition of glycine and then
265 allowing continued translation (Go). *RNA* **13**: 803–810.
- 266 Belsham GJ. 2005. Translation and replication of FMDV RNA. *Curr Top Microbiol Immunol*
267 **288**: 43–70.
- 268 Bulmer M. 1991. The selection-mutation-drift theory of synonymous codon usage. *Genetics*
269 **129**: 897–907.
- 270 Chan PP, Lowe TM. 2009. GtRNADB: A database of transfer RNA genes detected in
271 genomic sequence. *Nucl Acids Res* **37**:D93-D97.
- 272 Donnelly ML, Gani D, Flint M, Monaghan S, Ryan MD. 1997. The cleavage activities of
273 aphthovirus and cardiovirus 2A proteins. *J Gen Virol* **78**: 13–21.
- 274 Donnelly ML, Luke G, Mehrotra A, Li X, Hughes LE, Gani D, Ryan MD. 2001a. Analysis of
275 the aphthovirus 2A/2B polyprotein “cleavage” mechanism indicates not a proteolytic
276 reaction, but a novel translational effect: A putative ribosomal “skip.” *J Gen Virol* **82**:
277 1013–1025.
- 278 Donnelly ML, Luke GA, Hughes LE, Luke G, Mendoza H, Dam E, Gani D, Ryan MD.
279 2001b. The “cleavage” activities of foot-and-mouth disease virus 2A site-directed
280 mutants and naturally occurring “2A-like” sequences. *J Gen Virol* **82**: 1027–1041.
- 281 Doronina VA, Wu C, de Felipe P, Sachs MS, Ryan MD, Brown JD. 2008. Site-specific
282 release of nascent chains from ribosomes at a sense codon. *Mol Cell Biol* **28**: 4227–
283 4239.
- 284 Ellard FM, Drew J, Blakemore WE, Stuart DI, King AMQ. 1999. Evidence for the role of His
285 142 of protein 1C in the acid induced disassembly of foot and mouth disease virus
286 capsids. *J Gen Virol* **80**: 1911–1918.
- 287 Gao ZL, Zhou JH, Zhang J, Ding YZ, Liu YS. 2014. The silent point mutations at the

- 288 cleavage site of 2A/2B have no effect on the self-cleavage activity of 2A of foot-and-
289 mouth disease virus. *Infect Genet Evol* **28**: 101–106.
- 290 Gullberg M, Polacek C, Bøtner A, Belsham GJ. 2013. Processing of the VP1/2A junction is
291 not necessary for production of foot-and-mouth disease virus empty capsids and
292 infectious viruses: characterization of “self-tagged” particles. *J Virol* **87**: 11591–11603.
- 293 Hahn H, Palmenberg AC. 1996. Mutational analysis of the encephalomyocarditis virus
294 primary cleavage. *J Virol* **70**: 6870–6875.
- 295 Kristensen T, Normann P, Gullberg M, Fahnøe U, Polacek C, Rasmussen TB, Belsham GJ.
296 2017. Determinants of the VP1/2A junction cleavage by the 3C protease in foot-and-
297 mouth disease virus-infected cells. *J Gen Virol* **98**: 385–395.
- 298 Loughran G, Libbey JE, Uddowla S, Scallan MF, Ryan MD, Fujinami RS, Rieder E, Atkins
299 JF. 2013. Theiler’s murine encephalomyelitis virus contrasts with encephalomyocarditis
300 and foot-and-mouth disease viruses in its functional utilization of the StopGo non-
301 standard translation mechanism. *J Gen Virol* **94**: 348–353.
- 302 Martinez-Salas E, Belsham GJ. 2017. Genome Organisation, Translation and Replication of
303 Foot-and-mouth Disease Virus RNA. In *Foot-and-Mouth Disease: Current Research
304 and Emerging Trends* (eds. F. Sobrino and E. Domingo), pp. 13–42, Caister Academic
305 Press.
- 306 Mauro VP, Chappell SA. 2014. A critical analysis of codon optimization in human
307 therapeutics. *Trends Mol Med* **20**: 604–613.
- 308 Nayak A, Goodfellow IG, Belsham GJ. 2005. Factors required for the Uridylylation of the
309 foot-and-mouth disease virus 3B1, 3B2, and 3B3 peptides by the RNA-dependent RNA
310 polymerase (3Dpol) in vitro. *J Virol* **79**: 7698–7706.
- 311 Novoa EM, Ribas de Pouplana L. 2012. Speeding with control: Codon usage, tRNAs, and
312 ribosomes. *Trends Genet* **28**: 574–581.
- 313 Palmenberg AC, Parks GD, Hall D, Ingraham RH, Seng TW, Pallal P V. 1992. Proteolytic
314 processing of the cardioviral cleavage in clone-derived P2 region□: primary 2A / 2B
315 precursors. *Virology* **190**: 754–762.
- 316 Ryan MD, Donnelly M, Lewis A, Mehrotra AP, Wilkie J, Gani D. 1999. A Model for
317 Nonstoichiometric, Cotranslational Protein Scission in Eukaryotic Ribosomes. *Bioorg*

- 318 *Chem* **27**: 55–79.
- 319 Ryan MD, Drew J. 1994. Foot-and-mouth disease virus 2A oligopeptide mediated cleavage
320 of an artificial polyprotein. *EMBO J* **13**: 928–33.
- 321 Ryan MD, King AMQ, Thomas GP. 1991. Cleavage of foot-and-mouth disease virus
322 polyprotein is mediated by residues located within a 19 amino acid sequence. *J Gen*
323 *Viro* **72**: 2727–2732.
- 324 Sharma P, Yan F, Doronina VA, Escuin-Ordinas H, Ryan MD, Brown JD. 2012. 2A peptides
325 provide distinct solutions to driving stop-carry on translational recoding. *Nucleic Acids*
326 *Res* **40**: 3143–3151.
- 327 Tulloch F, Luke GA, Ryan MD. 2017. Foot-and-mouth disease virus proteinases and
328 polyprotein processing. In *Foot-and-Mouth Disease: Current Research and Emerging*
329 *Trends* (eds. F. Sobrino and E. Domingo), pp. 43–59, Caister Academic Press.
- 330
- 331
- 332
- 333
- 334
- 335

336 **Table 1: Codon utilization encoding the “NPGP” motif at the 2A/2B junction within**
 337 **rescued FMDVs.**

| Residue | Codon | | p2 % ¹ | p3 % ¹ | p4 % ¹ |
|---------|------------|----|----------------------|----------------------|----------------------|
| N16 | <i>AAT</i> | | 45 | 25 | 5 |
| N16 | <i>AAC</i> | wt | 55 | 75 | 95 |
| P17 | <i>CCT</i> | wt | 55 | 70 | 100 |
| P17 | <i>CCC</i> | | 10 | 15 | 0 |
| P17 | <i>CCA</i> | | 20 | 5 | 0 |
| P17 | <i>CCG</i> | | 15 | 10 | 0 |
| G18 | <i>GGT</i> | | 15 | 5 | 0 |
| G18 | <i>GGC</i> | | 15 | 5 | 0 |
| G18 | <i>GGA</i> | | 50 | 40 | 10 |
| G18 | <i>GGG</i> | wt | 20 | 50 | 90 |
| P19 | <i>CCT</i> | | 50 | 25 | 5 |
| P19 | <i>CCC</i> | wt | 10 | 50 | 95 |
| P19 | <i>CCA</i> | | 10 | 5 | 0 |
| P19 | <i>CCG</i> | | 30 | 20 | 0 |

338

339 **1:** From sequencing of plasmid DNA isolated from separate 20 colonies in each case, the
 340 proportion (%) of each codon present in the rescued FMDVs is indicated at the different
 341 passage (p) numbers. Codon frequency values of 50-70% are highlighted in light grey
 342 whereas values from 75-100% are highlighted in dark grey.

343

344

345 **Table 2: Enrichment for non-wt sequences encoding residues G¹⁸ and P¹⁹ within the**
 346 **“NPGP” motif within rescued FMDVs.**

| Residue | Codon | | Pretreatment | p2 % ² | p3 % ² |
|---------|------------|----|--------------------------|----------------------|----------------------|
| G18 | <i>GGT</i> | | <i>ApaI</i> ¹ | 5 | 0 |
| G18 | <i>GGC</i> | | <i>ApaI</i> ¹ | 15 | 5 |
| G18 | <i>GGA</i> | | <i>ApaI</i> ¹ | 80 | 95 |
| G18 | <i>GGG</i> | wt | <i>ApaI</i> ¹ | 0 | 0 |
| P19 | <i>CCT</i> | | <i>ApaI</i> ¹ | 60 | 55 |
| P19 | <i>CCC</i> | wt | <i>ApaI</i> ¹ | 0 | 0 |
| P19 | <i>CCA</i> | | <i>ApaI</i> ¹ | 15 | 0 |
| P19 | <i>CCG</i> | | <i>ApaI</i> ¹ | 25 | 45 |

347

348 **1:** Following RT-PCR, the 650bp amplicons were digested with *ApaI* to enrich the population
 349 in non-wt sequences and the residual intact products were inserted into the pCR-XL-TOPO
 350 vector (see text).

351 **2:** From sequencing of plasmid DNA isolated from separate 20 colonies in each case, the
 352 proportion (%) of each codon present in the rescued FMDVs is indicated at the different
 353 passage (p) numbers. Codon frequency values of 50-70% are highlighted in light grey
 354 whereas values from 75-100% are highlighted in dark grey.

355

356

357

358 **Table 3: Primers used to create and sequence mutant FMDV cDNAs.**

| Primer | Sequence (5'-3') |
|-------------------|--|
| Fwd_2A_N16A_degen | GGAGTCC <u>NNN</u> CCTGGGCCCTTC |
| Fwd_2A_P17A_degen | GTCCAAC <u>NNN</u> GGGCCCTTC |
| Fwd_2A_G18A_degen | GACGTCGAGTCCAACCCT <u>NNN</u> CCCTTCTTTTTCTCCGACGTTA |
| Fwd_2A_P19G_degen | TCG AGTCCAACCCTGGG <u>NNN</u> TTCTTTTTCTCCGACGTTAGG |
| 8APN206 | CACCCGAAGACCTTGAGAG |
| 8APN203 | CTCCTTCAACTACGGTGCC |

359

360

361 **Figure legend.**

362

363 **Figure 1: Structure of the FMDV O1 Kaufbeuren (O1K) cDNA and its derivatives.** (A)

364 The plasmid-encoded amino acids and the corresponding nucleotide sequences at the 2A/2B

365 junction are shown. The FMDV O1K degenerate codon mutants were produced as described

366 in the text using the mutant pT7S3 plasmids encoding the N¹⁶A, P¹⁷A, G¹⁸A and P¹⁹A

367 substitutions as templates. The full-length plasmid pools were linearized using *HpaI* prior to

368 *in vitro* transcription and virus rescue. The locations of the *HpaI* and *ApaI* restriction sites

369 that were used are marked. N = a mixture of the 4 nucleotides. (B) Chromatograms and

370 sequences of the FMDV cDNA corresponding to the NPGP motif at the 2A/2B junction.

371 Degenerate positions showing the presence of multiple nucleotides are marked with an N (in

372 bold type). The colour code in the chromatograms is as follows: A (red), T (green), G

373 (yellow), C (blue).

374

375

376

377

378

379

380

381

382

383

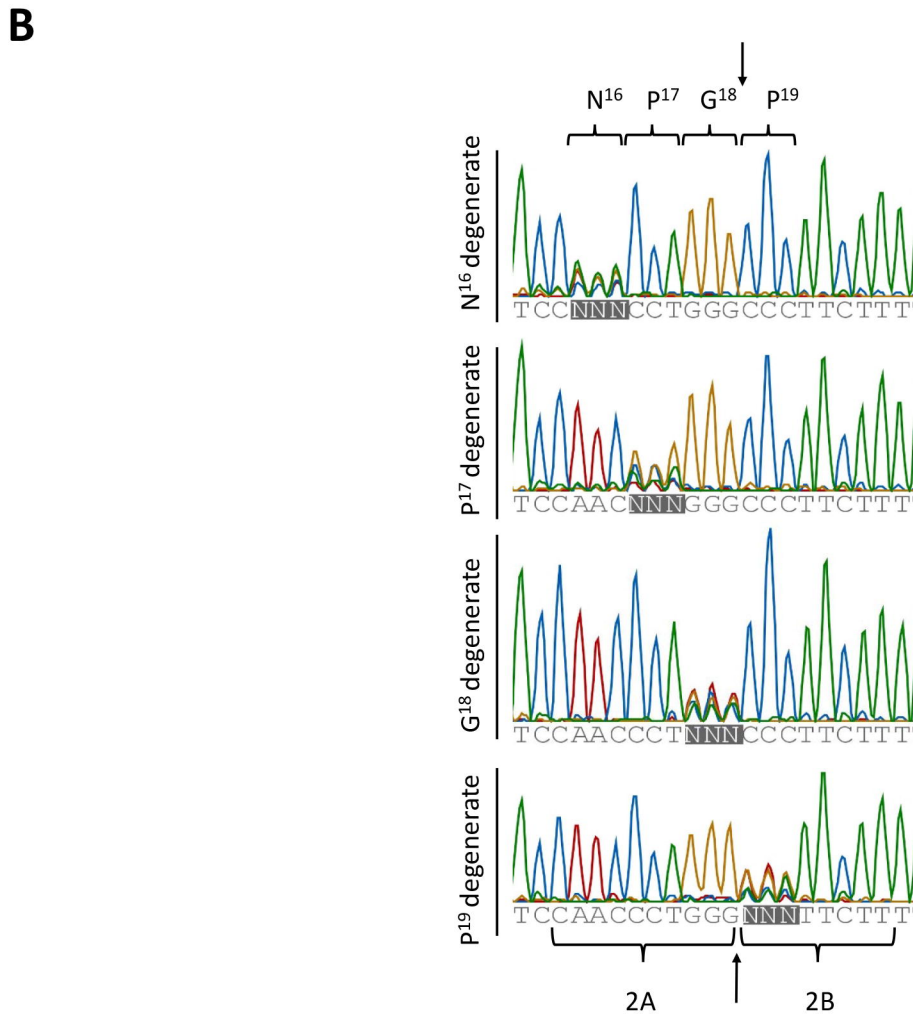
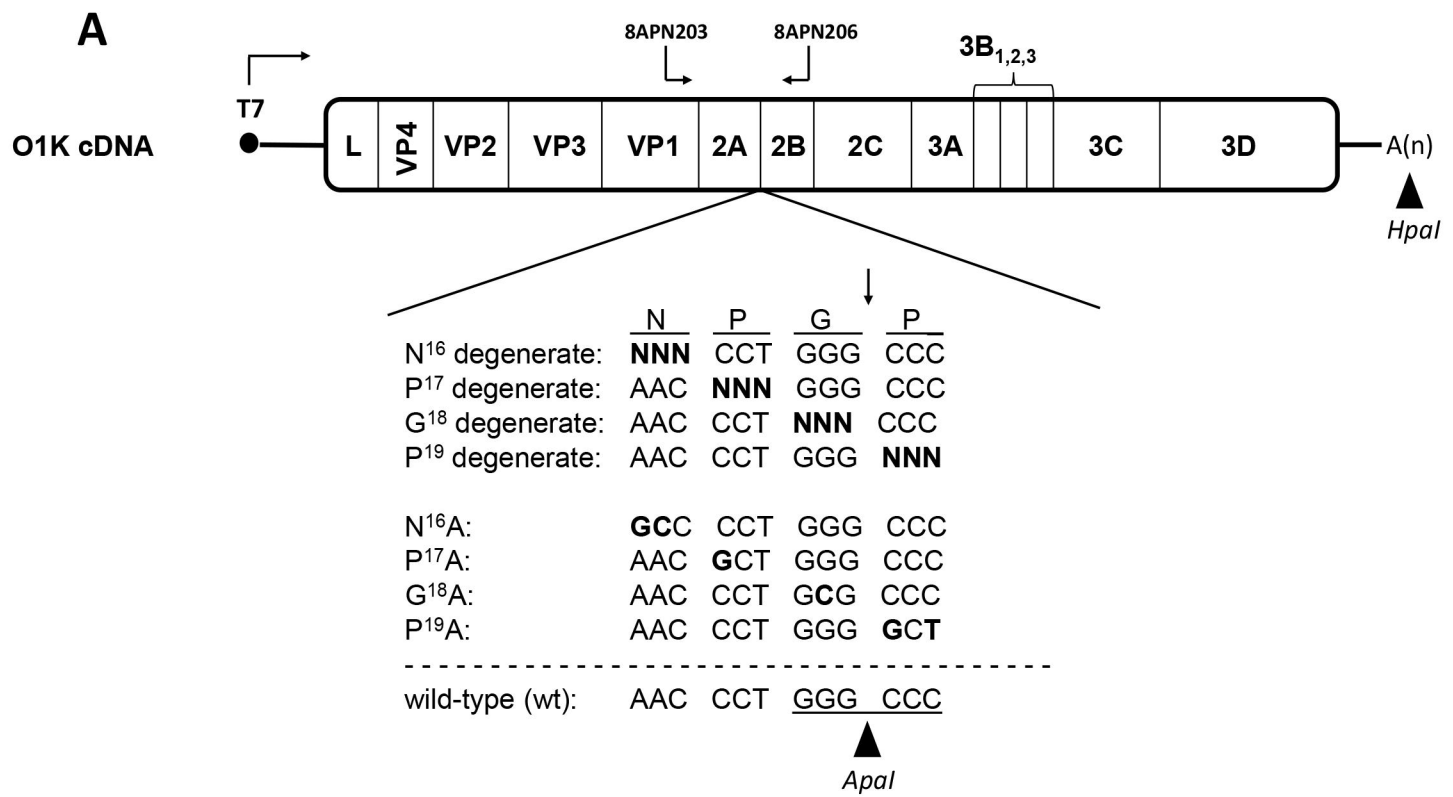
384

385

386

387

388





RNA

A PUBLICATION OF THE RNA SOCIETY

Selection of functional 2A sequences within foot-and-mouth disease virus; requirements for the NPGP motif with a distinct codon bias

Jonas Kjær and Graham J. Belsham

RNA published online October 17, 2017

P<P Published online October 17, 2017 in advance of the print journal.

Accepted Manuscript Peer-reviewed and accepted for publication but not copyedited or typeset; accepted manuscript is likely to differ from the final, published version.

Creative Commons License This article is distributed exclusively by the RNA Society for the first 12 months after the full-issue publication date (see <http://rnajournal.cshlp.org/site/misc/terms.xhtml>). After 12 months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).



To subscribe to *RNA* go to:
<http://rnajournal.cshlp.org/subscriptions>
