Published online 2 December 2011 *Nucleic Acids Research, 2012, Vol. 40, No. 7* ***3143–3151***
doi:10.1093/nar/gkr1176

2A peptides provide distinct solutions to driving stop-carry on translational recoding

Pamila Sharma[1], Fu Yan[1], Victoria A. Doronina[1], Helena Escuin-Ordinas[2], Martin D. Ryan[2] and Jeremy D. Brown[1,*]

[1]RNA Biology Group and Institute for Cell and Molecular Biosciences, The Medical School, Newcastle University, Newcastle upon Tyne NE2 4HH and [2]Centre for Biomolecular Sciences, Biomolecular Sciences Building, North Haugh, University of St. Andrews, St. Andrews KY16 9ST, UK

Received October 17, 2011; Revised November 10, 2011; Accepted November 11, 2011

ABSTRACT

Expression of viral proteins frequently includes non-canonical decoding events ('recoding') during translation. '2A' oligopeptides drive one such event, termed 'stop-carry on' recoding. Nascent 2A peptides interact with the ribosomal exit tunnel to dictate an unusual stop codon-independent termination of translation at the final Pro codon of 2A. Subsequently, translation 'reinitiates' on the same codon, two individual proteins being generated from one open reading frame. Many 2A peptides have been identified, and they have a conserved C-terminal motif. Little similarity is present in the N-terminal portions of these peptides, which might suggest that these amino acids are not important in the 2A reaction. However, mutagenesis indicates that identity of the amino acid at nearly all positions of a single 2A peptide is important for activity. Each 2A may then represent a specific solution for positioning the conserved C-terminus within the peptidyl-transferase centre to promote recoding. Nascent 2A peptide:ribosome interactions are suggested to alter ribosomal fine structure to discriminate against prolyl-tRNA^{Pro} and promote termination in the absence of a stop codon. Such structural modifications may account for our observation that replacement of the final Pro codon of 2A with any stop codon both stalls ribosome processivity and inhibits nascent chain release.

INTRODUCTION

A small number of nascent peptides are known to interact with the ribosomal exit tunnel and/or peptidyl-transferase centre (PTC) to regulate translation of downstream mRNA sequences (1,2). Most of these are encoded by short upstream ORFs (uORFs) in eukaryotes, or leader peptides in bacteria, and typically pause or stall ribosomes in response to small molecule effectors. '2A' peptides, characterized by a C-terminal '-D_{12}(V/I)ExNPGP$_{19}$-' motif (3,4), are further sequences that pause ribosomes (5–7). However, unlike other characterized ribosome-pausing peptides, 2As also drive a translational recoding event, in which two separate proteins (the 'upstream' and 'downstream' products) are generated from an ORF containing a 2A coding sequence, with a 'break' in the polypeptide backbone between the final 2 amino acids, Gly and Pro, of 2A (formally this Pro is the first amino acid of the downstream protein, but is an essential part of active 2A). The 2A reaction has been termed 'StopGo' or 'stop-carry on' recoding (6,8). 2A peptides are active when transposed into other proteins, i.e. they are autonomous elements, and mediate recoding in all eukaryotic ribosomes tested. The unique activity of 2A peptides has led to their use as tools for co-expression of two (or more) proteins in biomedicine and biotechnology (9).

The most studied 2A peptide is found in the *Aphthovirus* foot-and-mouth disease virus (FMDV). The FMDV RNA genome contains a single, long ORF encoding a polyprotein. 2A directs a key, primary processing event in the resolution of the polyprotein to individual proteins. Modelling of FMDV 2A indicated that it may form an (amphipathic) α-helix over most of its length, with a reverse turn at its C-terminus (5,10). This led to the suggestion that the α-helix, in the ribosomal exit

*To whom correspondence should be addressed. Tel: +44 0191 222 7470; Fax: +44 0191 222 7424; Email: jeremy.brown@ncl.ac.uk
Present addresses:
Pamila Sharma, University of Manchester, Manchester M13 9PT, UK.
Helena Escuin-Ordinas, Division of Hematology/Oncology, Department of Medicine, University of California Los Angeles, Los Angeles CA90095, USA.

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

tunnel, combined with the reverse turn, might subtly shift the peptidyl–tRNA^{Gly} in the PTC, disfavouring peptide bond formation to prolyl–tRNA^{Pro}, while promoting hydrolysis of the peptidyl–tRNA^{Gly} ester bond, thereby generating the upstream product. Pro is the poorest nucleophile of all proteinogenic amino acids since its secondary amino group is conformationally and sterically constrained. It may therefore be uniquely affected by subtle changes within the PTC. Hydrolysis of the nascent peptidyl-2A-tRNA^{Gly} ester bond is catalysed by translation terminating release factors (eRFs), despite the presence of the Pro codon in the A-site (6). Since mutation of this Pro codon to codons encoding several other amino acids almost completely abrogates 2A activity (5,11), any conformational shift in the PTC driven by 2A is suggested to discriminate prolyl-tRNA^{Pro} from other aminoacyl-tRNAs.

Sequencing of viral genomes has provided increasing numbers of 2A sequences and this provides opportunity to examine common features of and differences in this family of peptides. Here, viral 2A peptides are compared and FMDV 2A used as an exemplar for extended mutagenesis, providing in-depth analysis of a peptide that affects ribosomal processivity. A direct test of disturbance to ribosome conformation driven by 2A was provided by testing whether release factors could efficiently terminate translation when Pro19 was replaced with a stop codon.

MATERIALS AND METHODS

Strains and media

Yeast strains JDY4 (*ura3-Δ99*, *ade2-101*, *his3-Δ200*, *leu2-Δ1*, *trp1-Δ99*, cir^{o}, MAT**a**) and JDY705 (*ura3-52*, *ade2-101*, *his3-Δ200*, *leu2-Δ1*, *trp1-Δ63*, *ubr1Δ::LEU2*, MATα) were grown on media indicated.

2A sequences

A set of viral 2A peptides (4) supplemented with newly identified sequences (Supplementary Table S1) was used to examine amino acid composition and conservation using Excel (Microsoft) and Weblogo (12).

Constructs

[UBI-R2A-ADE2] was assembled in pRS314 (13) from fragments encoding ubiquitin, FMDV 2A preceded by arginine and Ade2p, [CFP-2A-PAC] in pcDNA3.1 (Invitrogen) from sequences encoding cyan fluorescent protein and puromycin acetyltransferase, each preceded by the V5 epitope tag and separated by FMDV 2A. Mutations were engineered into these constructs by replacing 2A sequences with PCR products, amplified from [UBI-R2A-ADE2], using oligonucleotides that incorporated the mutations. An ApaI site used in cloning corresponds to the final 2 codons of 2A. Mutations to these codons were introduced using fragments extended to include an ApaI site, thus inserting an extra GP dipeptide. Polymerase chain reaction (PCR) products from which mRNAs ending with 2A sequences were synthesized were generated from pJN141 (6) using SP6 primer and primers incorporating nucleotides of 2A up to the codon required and mutations as necessary. Clones were verified by sequencing.

Random mutagenesis

Mutations were incorporated into 2A amino acids Asp14 to Pro19 by amplifying a fragment from [UBI-R2A-ADE2] using oligonucleotides CM (5′-GACCTTCT TAAGCTTGCGGGA*GACGTCGAGTCCAACCCTGG GCCC*ATCGATCCCATG), which contained 2% of each non-template nucleotide at italicized positions, and adeR (5′-CATTAACGTGGTCATTGGAGTTGC). This was combined in a second round of PCR with an overlapping product extending from the T3 promoter upstream of [UBI-R2A-ADE2] to the 5′ portion of 2A (CMR 5′-TC CCGCAAGCTTAAGAAGGTC). The resulting product was transformed into JDY4 with the [UBI-R2A-ADE2] plasmid digested to remove the [UBI-R2A] coding region, and *in vivo* gap-repair regenerated complete plasmids incorporating mutations. A similar strategy introduced mutations into sequences coding for Leu1-Gly11. Oligonucleotides were NN (5′-TCTAGAA GGAGCATGCCAG*CTGTTGAATTTTGACCTTCTTA AGCTTGCGGGA*GACGTCGAGTCCAACCC) plus adeR and NR (5′ CTGGCATGCTCCTCTAGA) plus T3.

Yeast transformations were plated onto media lacking tryptophan and containing limited adenine as hemisulphate (24 μg/ml). All white or pigmented colonies were picked from regions of transformation plates for unbiased selection. Each was mixed into 150 μl of 30% v/v glycerol in a microtitre plate well and a 48-pin tool used to spot cells onto media. A region of the plasmid in each colony including 2A was amplified from yeast cells and sequenced.

In vitro transcription/translation, quantification

[CFP-2A-PAC] and derivatives were analysed in rabbit reticulocyte lysate-coupled transcription-translation reactions (Promega), and quantified (3). Capped mRNAs (Figure 5B and C) were generated using SP6 RNA polymerase (Promega), and translated in wheat germ lysate (Promega). Translation products were resolved on Tris–Tricine (14), or NuPAGE (Invitrogen) sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS–PAGE) gels, images captured and analysed on a Typhoon Trio scanner with ImageQuant (GE Healthcare).

RESULTS

Reporters of 2A activity

To facilitate analysis of 2A and mutant versions of it, two artificial polyprotein reporters were used. The first comprised cyan fluorescent protein-2A-puromycin acetyltransferase ([CFP-2A-PAC]) (Figure 1A) and templates encoding this were used in *in vitro* translation reactions with [35-S] methionine, from which the products were quantified after resolution on SDS–PAGE gels. The second permitted assessment of 2A activity *in vivo* in yeast (*Saccharomyces cerevisiae*), where we have

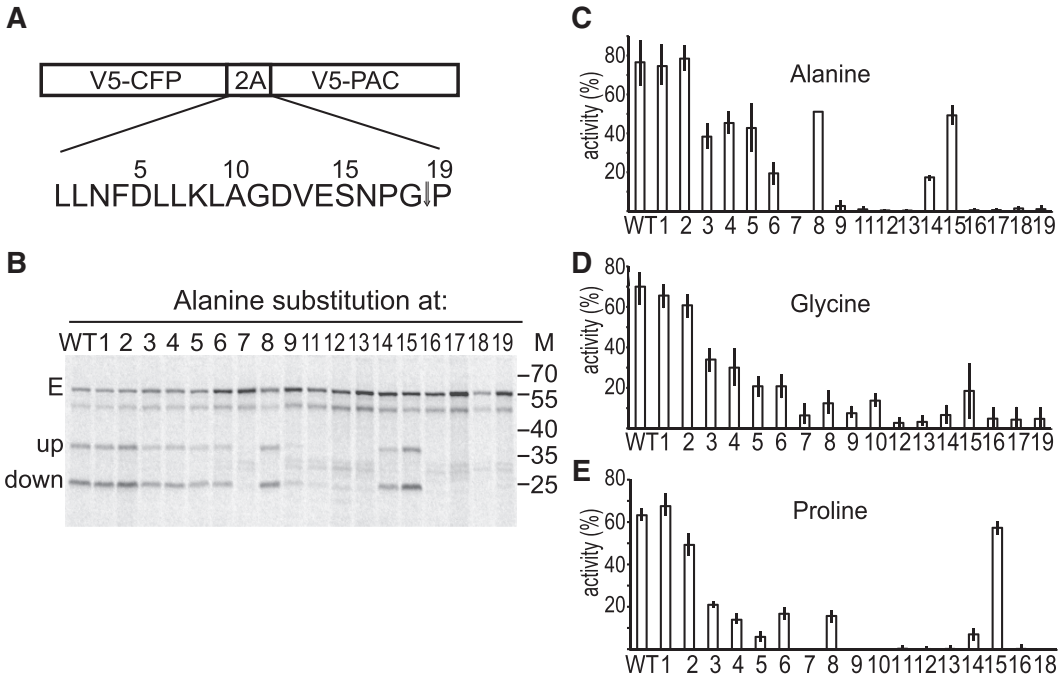


Figure 1. Scanning mutagenesis of FMDV 2A. (**A**) The [CFP-2A-PAC] reporter polyprotein comprises cyan fluorescent protein and puromycin acetyltransferase, each preceded by a single V5 epitope tag and separated by FMDV 2A. The separation point of the nascent chain driven by the 2A reaction is indicated with a downwards arrow. (**B–E**) Activities of [CFP-2A-PAC] variants containing (B, C) Ala (D) Gly or (E) Pro substituted for the individual amino acids of 2A indicated, and the wild-type (wt). (B) Coupled reticulocyte lysate transcription–translation reactions assembled with [35-S] methionine were programmed with constructs encoding [CFP-2A-PAC] or Ala substitutions at the positions indicated, resolved on an SDS–PAGE gel. M, markers in kDa. E, up and down refer to the complete translation product [CFP-2A-PAC], and the upstream [CFP-2A] and downstream [PAC] products of the 2A reaction, respectively. (C–E) Three independent sets of experiments as in (B) were quantified for each set of substitutions. Average activity (% of ribosomes releasing the upstream product) ± one standard deviation is shown.

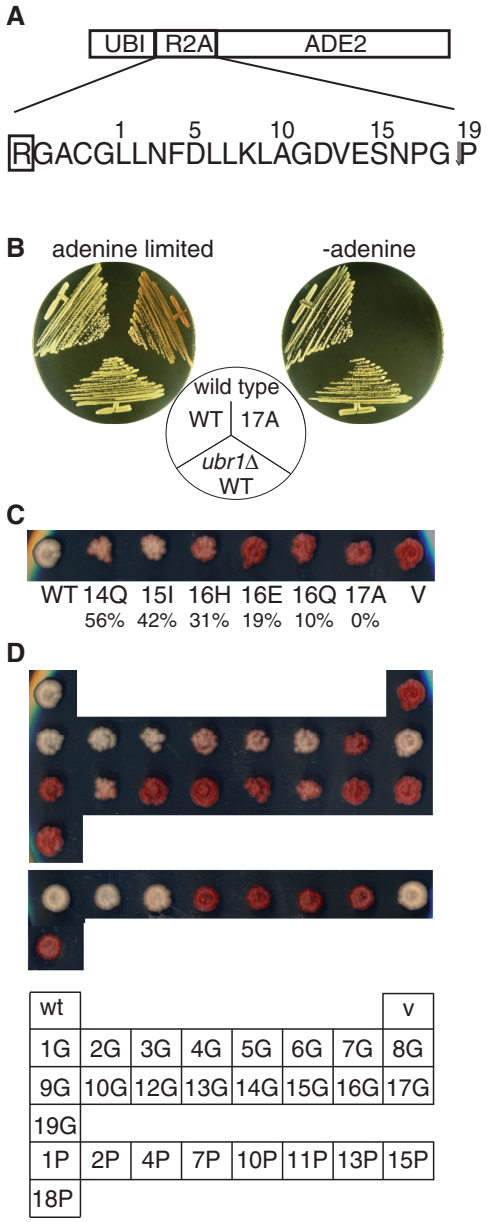
previously shown 2A to be active (6,15), and comprised ubiquitin–arginine-2A-phosphoribosylaminoimidazole (AIR) carboxylase ([UBI-R2A-ADE2]) (Figure 2A). Ubiquitin-dependent proteases cleave ubiquitin fusion proteins, and thus ubiquitin will be removed from [UBI-R2A-ADE2] leaving Arg as a new N-terminal amino acid. An N-terminal Arg targets a protein for degradation through the N-end rule pathway (16), and stable Ade2p is only generated from this reporter if the 2A reaction has freed it from the N-terminal Arg. Thus, expression of [UBI-R2A-ADE2] in otherwise *ade2* yeast leads to growth on plates lacking adenine, while expression of a variant containing the non-functional Pro17 to Ala 2A mutant (3) does not (Figure 2B). N-end rule pathway-deficient cells [lacking the E3-ubiquitin ligase Ubr1p (17)] expressing this mutant grow on media lacking adenine, confirming that [R2A-ADE2] is degraded through the N-end rule pathway.

Yeast cells lacking Ade2p accumulate red derivatives of the enzyme's substrate, AIR (18) and colonies of *ade2* cells expressing the Pro17 to Ala mutant of [UBI-R2A-ADE2] were pigmented on plates containing limited amounts of adenine (Figure 2B). Further, [UBI-R2A-ADE2] allowed discrimination between 2A mutants with different activities. Cells expressing [UBI-R2A-ADE2] variants containing previously characterized partially active 2A mutants (3) displayed a range of colours, the most active mutants being pale pink and those with less activity being more pigmented (Figure 2C). The relative activity of the mutants was slightly different in the context of [UBI-R2A-ADE2] than in *in vitro* translations [compare colony colours in Figure 2C with the % activity figures below, which are from (3)]. These differences may be due to translation by yeast, rather than wheat or rabbit ribosomes.

2A is largely intolerant to sequence alteration

Previous mutagenesis of 2A has largely been to conserved C-terminal residues, and individual positions within the peptide have not been substituted with a wide range of other amino acids. To examine sequence constraints within 2A, systematic Pro-, Gly- and Ala-scanning as well as random mutagenesis were carried out on FMDV 2A. Pro and Gly influence secondary structure, constraining or providing high flexibility to peptide chains, respectively. In contrast, Ala has little influence on secondary structure. Complete sets of mutants with Pro, Gly and Ala substituted for individual amino acids in 2A were generated in [CFP-2A-PAC], and analysed by *in vitro* translation (Figure 1B–E). The Gly and a subset of the Pro substitution mutants were also tested *in vivo* in the context of [UBI-R2A-ADE2].

Qualitatively the *in vitro* and *in vivo* assays provided similar results, Gly and, where tested, Pro substitution mutants showing the least activity *in vitro* yielding



wt							v
1G	2G	3G	4G	5G	6G	7G	8G
9G	10G	12G	13G	14G	15G	16G	17G
19G							
1P	2P	4P	7P	10P	11P	13P	15P
18P							

Figure 2. The [UBI-R2A-ADE2] artificial polyprotein reporter. (**A**) [UBI-R2A-ADE2] comprises ubiquitin and *S. cerevisiae* Ade2p separated by a single arginine, a 4 amino acid linker (GACG) and FMDV 2A. The separation point of the nascent chain driven by the 2A reaction is indicated with a downwards arrow. (**B**) [UBI-R2A-ADE2] containing wild-type 2A or the non-functional Pro17 to Ala mutant (3) were expressed in *ade2* yeast strains that were either contained (wild type), or lacked (*ubr1Δ*), and streaked on plates containing limited (left) or no (right) adenine as indicated (**C**) *ade2* yeast expressing [UBI-R2A-ADE2] containing 2A mutants as indicated were spotted onto plates containing limited amounts of adenine. The previously determined activity of each mutant in *in vitro* translations (3) is indicated. (**D**) *ade2* yeast expressing [UBI-2A-ADE2], Gly- or Pro-substitution mutants or an empty vector as indicated in the table below were assayed by spotting suspensions of the cells onto plates containing limited amounts of adenine. Images of plates were captured using an Epson scanner, imported into and cropped in Photoshop. In (B–D) wild-type 2A is designated WT, and mutants by the numerical position in 2A and the single letter code for the amino acid replacing the wild type at that position. V is an empty vector control. Plates were incubated for 3 days at 30°C.

colonies with the highest degree of pigmentation (compare Figure 1D and E with Figure 2D). Substitutions at most positions of 2A for Ala, Gly or Pro had similar effects, and, in most cases, activity was more affected by substitutions to amino acids towards the C-terminus of 2A than those within its N-terminal portion. Only substitutions of Leu1 or Leu2 were neutral, substitutions at positions 3–6 reduced activity between a half to one-third of wild type (though Pro at position 5 had very low activity), while substitutions to amino acids 7–19 had, in general, low activity. Exceptions were Ala at positions 8 and 15, which were around 50% active *in vitro* and Pro at position 15, which yielded nearly 100% activity *in vitro* and persistently white colonies *in vivo*. Ala is not found at position 8 or 15 in known viral 2A peptides, although Pro is observed at residue 15 in Duck hepatitis virus (DHV-1) and in a small number of FMDV strains (Figure 4A; Supplementary Table S1 and data not shown). Replacement of position 15 with Gly, in contrast to Ala or Pro, yielded low activity. Flexibility provided to the nascent chain by introduction of Gly at this position may adversely affect its ability to adopt a conformation conducive to the 2A reaction.

Screens for random mutants within 2A were performed using the [UBI-R2A-ADE2] reporter polyprotein (see 'Materials and Methods' section). This provided 52 unique single amino acid alterations to 2A, and this collection was supplemented with site-directed mutants that provided alterations at positions where no mutants had been generated in the screens, and also broadened the spectrum of substitutions represented at some other positions. Including mutants generated in the scanning mutagenesis, 92 single amino acid changes to 2A were created within [UBI-R2A-ADE2]. Mutants were assigned to four activity categories: white (active), pale pink, dark pink and red (inactive) (summarized in Figure 3). Only 15 mutants generated sufficiently high Ade2p activity to yield white colonies. Of these, eight were isolated from the random screens in which an equal number (200) of white and pigmented colonies were examined, and 11 were alterations to positions 1, 2 or 15. Mutations that substantially decreased 2A activity were identified at nearly all other positions of 2A, and included several conservative changes (e.g. Leu7 and Leu9 to Val). The exception was position 8, where all of the five changes examined yielded white or pale pink colonies. The only change to a conserved position of 2A that retained relatively high activity was Gln14 to Glu, which yielded pale pink colonies, and this confirmed previous results (3).

Comparison of 2A peptide sequences

Previous analysis of 2A peptides has indicated that they comprise a conserved C-terminal motif, -D_{12}(V/I) ExNPG⇓P_{19}- (3,4) and a non-conserved N-terminal portion. The finding that many changes to non-conserved as well as conserved amino acids within 2A led to significant loss of activity (Figures 1–3) was then somewhat surprising, given that conservation typically correlates with functional importance. The number of 2A peptides identified in viral genomes has increased in recent years

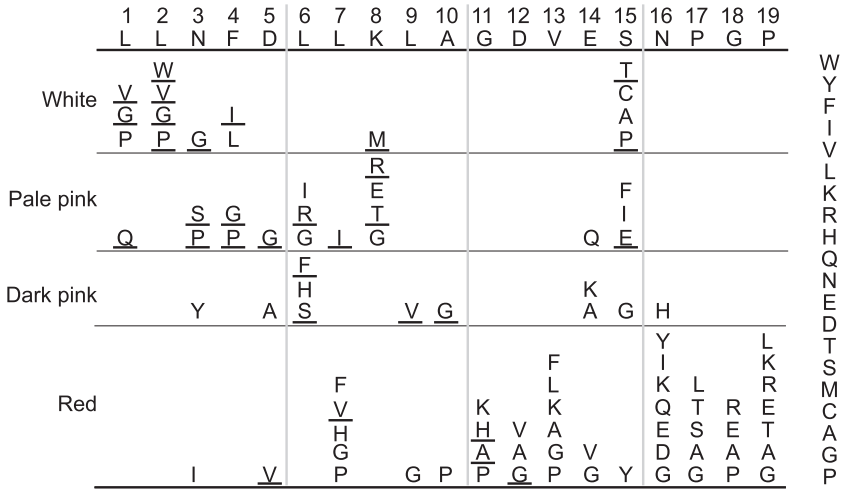


Figure 3. 2A is intolerant to mutation at most positions. The FMDV 2A sequence is shown, with, underneath, single amino acid changes isolated at each position in random screens or generated through site-directed mutagenesis and which were assessed *in vivo* using [UBI-R2A-ADE2]. Changes that correspond to amino acids found in other 2A peptides are underlined. Activity was assessed from colour of colonies after growth on plates containing limiting adenine (as in Figure 2) and divided into four categories: white (as with wild-type 2A), pale pink, dark pink and red (as cells lacking *ADE2* or expressing the Pro17 to Ala mutant of [UBI-R2A-ADE2]). Within each category, the order of amino acids does not reflect relative activity, but is as they appear in the list to the right.

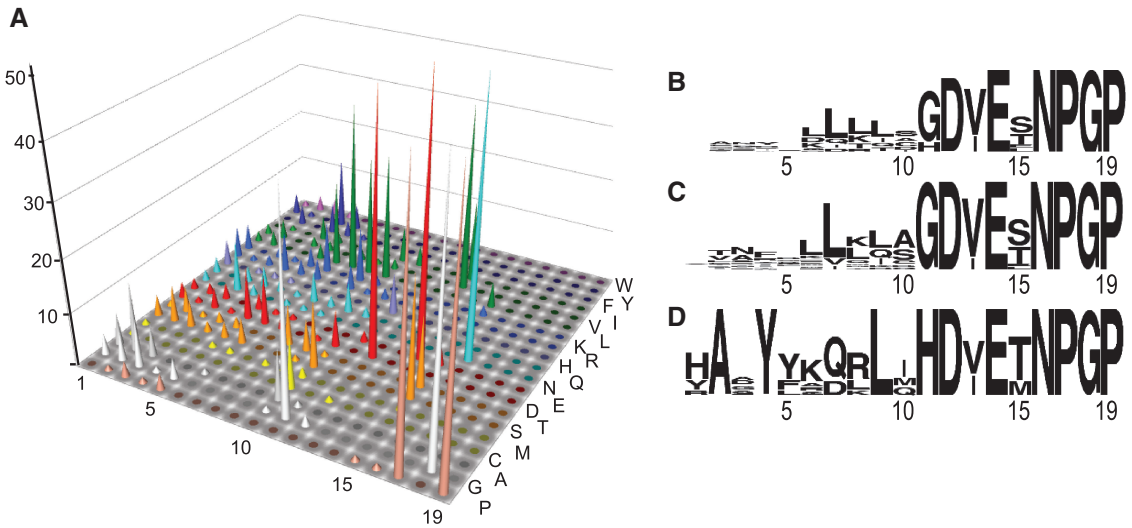


Figure 4. Sequence variation within 2A peptides. (**A**) Amino acids present at each position within 2A are shown in a cone plot, with height of cones corresponding to the number of times each amino acid is present at this position in the 52 2A peptides in the data set (sequences are in Supplementary Table S1). Amino acids are coloured according to the standard Rasmol amino colour scheme, and grouped by chemical property. (**B–D**) Conservation within 2A peptides is shown as Weblogo representations. Each stack of letters corresponds to one position within the sequence, with the height of each stack indicating the conservation at that position and the height of letters within the stack the frequency at which the corresponding amino acid is found. Separate Weblogos are shown for (B) the full set of 52 2A peptides (C) Picornaviral peptides with Gly at position 11 (17 sequences) and (D) Picornaviral peptides with His at position 11 (eight sequences).

and a comparison of the full set of 2A peptides available was made to determine whether this might aid interpretation of the mutagenesis data (Figure 4A and B and Supplementary Table S1). It should be noted that while the activity of many of these have been confirmed in reporter proteins (indicated in Supplementary Table S1), others have not been tested, and are thus currently putative 2As. This analysis confirmed the previous consensus, and highlighted position 11 as also being conserved. Gly is present at position 11 in 41 of the 52 sequences, and His in all but three of the remainder. Within the N-terminal portion of 2As, Leu is the most frequent amino acid at positions 6–9, while Gly, Pro and aromatic amino acids are excluded from positions 6–10 [a single exception being Phe6 in Ljungan virus (LV)]. The first five amino acids of 2A are extremely variable. However, these, and amino acids upstream of 2A itself, enhance reaction efficiency (3,4,19). Features, not readily identifiable from the primary sequence, must then be included in (or excluded from) this region for it to stimulate 2A function.

Integrated function of the N-terminal portion of 2A

Amongst the mutations that inactivated FMDV 2A were all those at position 11, including substitution with His (Figure 4). This was intriguing because, as noted above, His is the most common amino acid at this position after Gly. His11 is found almost exclusively in cardioviral 2A peptides, the sole exception being a human cosaviral 2A (HCoSV-B1). Comparison of this sub-set of sequences with other Picornaviral 2As (containing Gly11) revealed a very different composition in the N-terminal portions (compare Figure 4C and D and see Supplementary Table S1). Thus, Ala2, Tyr4 and Leu9 are present in all available His11-containing 2A sequences. In addition, Ile, Met or Gln is found at position 10 and Gln or Asp is only found at positions 7 of His11-containing, but not in Gly11-containing, 2A peptides. This suggests that very different constraints may exist on 2A sequence when His, rather than Gly, is present at position 11. More globally, this suggests that the sequence of the non-conserved portion of each distinct 2A may be a specific solution to the problem of allowing the conserved motif to drive the 2A reaction, rather than particular amino acid(s) being possible at each position independent of others.

Inactivity of the FMDV 2A containing His11 provided a means to test this hypothesis. EMCV and TMEV, which both contain His11, are active in [CFP-2A-PAC] and other artificial polyproteins [Table 1; (3,19)]. However, fusions between sequences encoding FMDV 2A and either EMCV or TMEV 2A peptides that altered position 11 with respect to upstream amino acids had similar, low, activity to the FMDV 2A His11 mutant (Table 1, mutants E9/10F, F9/10E, T9/10F, T10/11F and F9/10T).

In further tests, changes were made to FMDV 2A alongside the Gly11 to His mutation, introducing either single or groups of amino acids found in His11-containing 2A peptides (Table 1). None of the mutations tested, including replacing amino acids 4–7 with the sequence YYKQ (mutant $FMDV_{(F4Y,D5Y,L6K,L7Q,G11H)}$) found in the majority of His11 containing 2A peptides (Supplementary Table S1 and Figure 4D), produced significant activity in combination with His at position 11. Indeed, many of the additional mutations further decreased the activity of the FMDV 2A Gly11 to His mutant. The mutations tested were not an exhaustive set, but these data suggest that the negative effect of changing position 11 from Gly to His within FMDV 2A cannot be compensated for by other small changes within the peptide, and is consistent with the proposal that specific combinations of amino acids throughout the N-terminal portion of 2A enable it to function.

Table 1. 2A peptides are fine-tuned to function as a whole

2A peptide/ mutant	Sequence 5 10 15 19	Activity (%)	SD (±)
FMDV	LLNFDLLKLAGDVESNPGP	73.5	9
EMCV	YAGYFADLLIHDIETNPGP	94.0	0.74
TMEV	HADYYKQRLIHDVETNPGP	60.7	6.8
E9/10F	YAGYFADLLAGDVESNPGP	15.0	2.9
F9/10E	LLNFDLLKLIHDIETNPGP	18.2	4.4
T9/10F	HADYYKQRLAGDVESNPGP	9.9	2.7
T10/11F	HADYYKQRLIGDVESNPGP	11.9	3.3
F9/10T	LLNFDLLKLIHDVETNPGP	14.5	1.1
FMDV(G11H)	LLNFDLLKLAHDVESNPGP	9.1	2.7
FMDV(A10I)	LLNFDLLKLIGDVESNPGP	5.5	7.3
FMDV(A10I,G11H)	LLNFDLLKLIHDVESNPGP	6.2	2.8
FMDV(L6K)	LLNFDKLKLAGDVESNPGP	33.2	2.7
FMDV(L6K,G11H)	LLNFDKLKLAHDVESNPGP	1.3	2.2
FMDV(L7Q)	LLNFDLQKLAGDVESNPGP	2.5	1.8
FMDV(L7Q,G11H)	LLNFDLQKLAHDVESNPGP	3.1	0.9
FMDV(L7D)	LLNFDLDKLAGDVESNPGP	0.9	1.5
FMDV(L7D,G11H)	LLNFDLDKLAHDVESNPGP	3.1	3.1
FMDV(L6K,L7Q,G11H)	LLNFDKQKLAHDVESNPGP	1.2	2.4
FMDV(F4Y,D5Y,L6K, L7Q,G11H)	LLNYYKQKLAHDVESNPGP	0.2	0.3

Coupled transcription–translation reactions were carried out and analysed as in Figure 1 for the 2A peptides and mutants indicated. Activity is the percentage of ribosomes releasing the upstream product from three or more independent samples and ± 1 SD is shown. EMCV sequence is highlighted in pale grey, TMEV in dark grey. Fusions between peptides are denoted by the first letter of the virus from which each portion originates, with numbers between them corresponding to the fusion junction.

2A affects ribosomal ability to carry out termination at a stop codon

Nascent 2A peptide-exit tunnel interactions promote translation termination over Gly18-Pro19 peptide bond formation as the first step in the 2A reaction (6). As an extension of the analysis of (sense) mutations to 2A, a variant of [CFP-2A-PAC] containing a stop codon in place of Pro19 was generated and analysed *in vitro*. This tested whether release factors could interpret a stop codon and catalyse termination in the context of the peptidyl 2A–ribosome complex. This revealed, on neutral pH SDS–PAGE gels (NuPage; Invitrogen) that retain peptidyl–tRNA ester bonds, a species ~20 kDa larger than that of the [CFP-2A] product. Consistent with this representing a peptidyl–tRNA adduct, its mobility increased to the expected size on treatment with RNAse (Figure 5A). A similar-sized, RNAse-sensitive product was also seen, though faintly, in translation reactions programmed with RNA containing a Gly18 to stop mutant, but not reactions programmed with a Pro17 to stop mutant nor a Pro17 to Ala, Pro19 to stop double mutant.

A second set of templates were used to confirm this result (Figure 5B). In these, sequences upstream of those encoding 2A were derived from yeast alpha factor rather than CFP, and they were truncated to leave sequences encoding 2A at their 3′ ends. Translation of truncated RNAs without stop codons leads to ribosome stalling, with nascent chains remaining tRNA associated. This was seen for the RNA that extended to the end of 2A, but included the inactivating Pro17 to Ala mutation, an RNAse-sensitive product larger than the free polypeptide accumulating in translation reactions programmed with this RNA. In contrast, and as seen previously (6), on translation of a template ending in wild-type 2A, the nascent peptide was released from tRNA due to the 2A reaction. When the RNA ended at position 17 of 2A, but

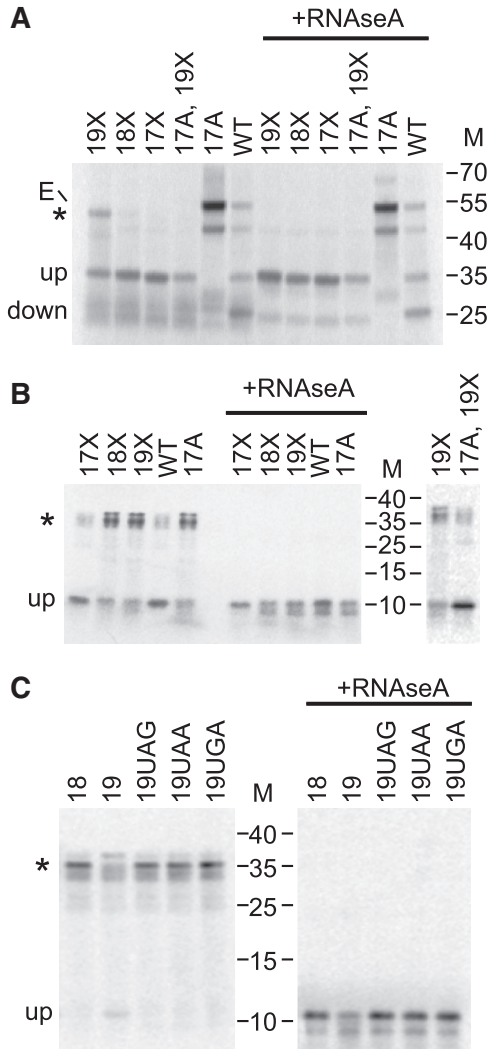


Figure 5. Replacement of G18 or P19 in FMDV 2A with a stop codon stalls ribosomes and impairs yeast growth. (**A**) Coupled *in vitro* transcription–translation reactions assembled in reticulocyte lysate with [35-S] methionine and indicated variants of [CFP-2A-PAC] were resolved on a Nu-PAGE (Invitrogen) SDS–PAGE gel and visualized using a phosphorimager. Full-length translation product (E), separated upstream, [CFP-2A] (up) and downstream PAC (down) products and the RNAse-sensitive peptidyl 2A-tRNA adduct (asterisks) are indicated. X indicates a stop codon (**B** and **C**) *In vitro* translation reactions assembled in wheat germ lysate with [35-S] methionine and programmed with truncated RNAs encoding abbreviated yeast pro-α-factor followed by 2A were analysed and labelled as in (A), except that here 'up' refers to [αF-2A] and in these experiments the RNA ended at the last codon indicated. Nucleotides of the three stop codons used to replace the Pro19 codon in (C) are indicated. Except for the two lanes in the right hand panel of (B), all samples were divided into two equal portions and one treated with RNAse A as indicated prior to loading. M, markers in kDa.

with the Pro17 codon replaced with a stop codon, the peptide was also efficiently released from tRNA, the stop codon functioning normally. However, translation of templates extending to position 18 or 19 of 2A, but with the Gly18 or Pro19 codon in each case replaced by a stop codon, led to strong accumulation of RNAse-sensitive peptidyl-tRNA adducts. This occurred on translation of templates ending at position 19 irrespective of which of the three stop codons was substituted for Pro19 (Figure 5C). Failure to release the nascent peptide from templates containing a stop codon instead of Pro19 was due to the presence of the otherwise wild-type 2A peptide sequence since the peptide was released on translation of a template containing both the Pro17 to Ala and Pro19 to stop mutations.

Data obtained from *in vitro* translation of the two sets of templates are consistent with a stop codon being poorly functional at position 19 of 2A, and the Gly18 to stop mutant peptide also being inefficiently released from ribosomes, particularly in the absence of further 3′ nucleotides. This provides evidence for 2A altering ribosomal structure and/or function, and further indicates that this is (partly) manifested one elongation cycle ahead of the position at which the 2A reaction takes place.

DISCUSSION

The 2A reaction co-translationally resolves what would otherwise be a continuous polypeptide chain into two, between the final Gly and Pro amino acids of 2A. The presence of functional 2A peptides in a wide variety of virus polyproteins indicates that they contribute significantly to viral gene expression. This is exemplified in EMCV, where inactivation of 2A results in loss of viability (11). Here, variability in naturally occurring viral 2A peptide sequences has been examined, and systematic mutagenesis of the FMDV 2A peptide performed. This was complemented with random screens for mutants and generation of fusions between 2A peptides with differing sequence compositions. Further, it was found that the 2A–ribosome interaction, in addition to preventing addition of Pro19 to the nascent chain, impedes the ability of eRFs to release the nascent chain when a stop replaced either Pro19 or Gly18 codon. This contrasts starkly with the normal 2A reaction in which eRFs release the nascent chain, and provides further strong evidence for 2A driving alteration to ribosomal conformation.

Mutations to 2A may increase or decrease the efficiency of the 2A reaction, alter the ratio of upstream to downstream product or, as seen with the Pro19 to stop mutants, impair release of the upstream product. FMDV 2A was ~70% efficient in the context of [CFP-2A-PAC], and thus increased reaction efficiency, as seen with EMCV 2A, could be revealed with this reporter (Table 1). However, none of the FMDV 2A mutants generated and examined in [CFP-2A-PAC] in this study increased its efficiency. Neither did any lead to significant alteration of the [CFP-2A]:PAC translation product ratio.

The [UBI-R2A-ADE2] reporter (Figure 2) facilitated rapid screening of randomly mutated 2A for effects on its activity. With regard to quantification of translational outcomes, the *in vivo* assay possessed a reduced dynamic range compared to *in vitro* translation. Thus, 2A mutants that yielded 50% or greater activity *in vitro* also yielded white colonies *in vivo* (e.g. Ala at position 15, Gly at positions 1 or 2). The amount of (stable) Ade2p generated from these mutants was still sufficient to saturate

requirement for this enzyme. At the other end of the activity scale, mutants with 20% activity or less in [CFP-2A-PAC] yielded red colonies with [UBI-R2A-ADE2]. In very few cases, the two assays gave differing readouts. The Asn3 to Gly mutant yielded white colonies, but only 35% activity *in vitro*, while Glu5 to Ala yielded dark pink colonies yet had ~45% activity *in vitro*. These differences possibly relate to differences in the ribosomes (rabbit or yeast) synthesizing the peptide. A limitation to [UBI-R2A-ADE2] is that it does not discriminate complete failure of the 2A reaction from decrease in synthesis of the downstream product. However, as noted above, no mutants were identified out of those tested in the [CFP-2A-PAC] reporter that altered the ratio of [CFP-2A]:PAC. Such mutations may then be rare.

The major conclusion from the FMDV 2A mutagenesis experiments is that it is largely intolerant to amino acid substitution over most of its length. This is highlighted by the fact that in the *in vivo* screen for 2A mutants only 8 of 200 white colonies picked contained mutations that altered the peptide sequence. While mutations of conserved amino acids have, in general, more pronounced effects than changes to non-conserved ones [Figures 1–3; (3)], alteration to several non-conserved positions (and particularly 7, 9 and 10) also yielded significant reduction, and in some cases complete loss of, activity. Substitutions at only two internal positions, 8 and 15, retained high activity (Figure 3). This differs to that observed in other peptides known to affect ribosomal activity such as secM, ermC and the arginine attenuator peptide, where mutations to a small number of critical amino acids have significant effects on activity (20–22). This supports the conclusion that the FMDV 2A, and by inference other 2A peptides, is optimized over its entire length to deliver the exacting requirements for translational recoding. Failure to generate 'hybrid' peptides with significant activity when the N- and C-terminal portions of FMDV, TMEV and EMCV 2A were swapped highlights the functional integration of each peptide (Table 1). The experiments carried out here tested 2A recoding activity, but it is worth noting that, at least for some 2A peptides, there may be other constraints on sequence in the context of the virus. This is highlighted by the identification of a mutation to the conserved portion of EMCV 2A that retains activity, but impedes subsequent proteolytic processing (11).

Picornaviral 2A peptides have been suggested to adopt an α-helical conformation within the exit tunnel over much of their length, with a reverse turn at the C-terminus (5,10,23). Although not necessarily informative in terms of the environment of the ribosomal exit tunnel, modelling of 2A peptides reveals that high α-helical propensity is not a universal feature within the 2A peptides examined here (data not shown). The different properties of Ala, Gly and Pro made it reasonable to expect that substitution of individual 2A amino acids within them may have different effects. However, these substitutions had, in general, similar effects, suggesting amino acid identity at most positions within the peptide is of primary importance, rather than any structural changes that the amino acid substitutions may impose on the peptide. An exception was position 15, where mutation of Ser15 to Gly impaired function more than Ala or Pro substitutions, suggesting that increased backbone flexibility conferred by Gly at this position was particularly detrimental. Until structural data of ribosomes with 2A peptides *in situ* are available, the conformation of the complex remains obscure. Structural data are, however, available for a number of other ribosome–nascent peptide complexes, both prokaryotic and eukaryotic, including some containing 'stalling' peptides (24–26). Several of these structures reveal stalling peptides in extended conformation in the exit tunnel. A cryo-EM structure of ribosomes containing the arginine attenuator peptide contains density close to the PTC consistent with a compact conformation of this portion of the peptide (25).

Mutants examined here increase the range of amino acid substitutions made to P19, the list extending to Gly, Leu, Lys, The, Ala, Glu, Arg, Ser, Ile and Phe Figures 2 and 3 (3,10)]. All these lead to loss or almost complete loss of activity. While not an exhaustive list, this supports the notion that the constrained amino group within the pyrrolidine ring of proline is key to the 2A reaction, and that ribosomal conformation is subtly altered by interactions with the nascent 2A peptide to specifically inhibit its incorporation into the nascent chain (5,27).

Further support for alteration to ribosomal conformation being at the heart of the 2A reaction comes from the remarkable observation that a stop codon at position 19 functions very poorly (Figure 4A and B). The inability to release the nascent chain in 2A containing a stop codon at position 19 contrasts starkly with translation of wild-type 2A where eRFs catalyse release of the nascent chain on the Pro19 codon. An interpretation of these data is that, analogous to the inability of prolyl-tRNA^{Pro} to decode a proline codon and incorporate proline into the nascent chain in wild type 2A, eRF1 recognizes the stop codon at position 19, but is then unable to engage productively with the PTC to catalyse hydrolysis of the peptidyl–tRNA^{Gly} ester linkage.

2A initiates changes to ribosomal conformation before the Pro19 codon reaches the ribosomal A-site, since templates including the Gly18 to stop mutant also lead to accumulation of peptidyl(2A)–tRNA adducts when the RNA ended at this point (Figure 4A). Since such adducts were seen at a much lower level when the RNA extended beyond 2A, mRNA beyond the A-site within the entry channel may oppose the effect of the 2A nascent chain in the exit tunnel until the Pro19 codon reaches the A-site.

In addition to 2A, stalling of ribosomes with a stop codon in the PTC is driven by several other peptides including the bacterial TnaC leader, cytomegalovirus UL4 gene uORF2 and fungal AAP. In each of the TnaC–uORF2- and AAP–ribosome complexes, the interaction of the peptide with the ribosome results in changes to the conformation of the PTC that are suggested to prevent release factor activity (24,25). Given that wild-type 2A 'promotes' release factor action, an expectation would be that the PTC of ribosomes at 2A is not distorted to impede termination activity.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Table 1 and Supplementary References (28 and 29).

ACKNOWLEDGEMENTS

The authors thank Nils Johnsson for JDY705, Nick Watkins and laboratory members for constructive discussions and/or comments on the manuscript and Mike Hoffman for Bluegill Picornavirus 2A sequences.

FUNDING

The U.K. Biotechnology and Biological Sciences Research Council (BB/E009093 to J.D.B., BB/E/01070911 to M.D.R.); Universities UK and Newcastle University (to P.S.). Funding for open access charge: BBSRC.

Conflict of interest statement. None declared.

REFERENCES

1. Ito,K., Chiba,S. and Pogliano,K. (2010) Divergent stalling sequences sense and control cellular physiology. *Biochem. Biophys. Res. Commun.*, **393**, 1–5.
2. Cruz-Vera,L.R., Sachs,M.S., Squires,C.L. and Yanofsky,C. (2011) Nascent polypeptide sequences that influence ribosome function. *Curr. Opin. Microbiol.*, **14**, 160–166.
3. Donnelly,M.L., Hughes,L.E., Luke,G., Mendoza,H., ten Dam,E., Gani,D. and Ryan,M.D. (2001) The 'cleavage' activities of foot-and-mouth disease virus 2A site-directed mutants and naturally occurring '2A-like' sequences. *J. Gen. Virol.*, **82**, 1027–1041.
4. Luke,G.A., de Felipe,P., Lukashev,A., Kallioinen,S.E., Bruno,E.A. and Ryan,M.D. (2008) Occurrence, function and evolutionary origins of '2A-like' sequences in virus genomes. *J. Gen. Virol.*, **89**, 1036–1042.
5. Donnelly,M.L., Luke,G., Mehrotra,A., Li,X., Hughes,L.E., Gani,D. and Ryan,M.D. (2001) Analysis of the aphthovirus 2A/2B polyprotein 'cleavage' mechanism indicates not a proteolytic reaction, but a novel translational effect: a putative ribosomal 'skip'. *J. Gen. Virol.*, **82**, 1013–1025.
6. Doronina,V.A., Wu,C., de Felipe,P., Sachs,M.S., Ryan,M.D. and Brown,J.D. (2008) Site-specific release of nascent chains from ribosomes at a sense codon. *Mol. Cell. Biol.*, **28**, 4227–4239.
7. Yan,F., Doronina,V.A., Sharma,P. and Brown,J.D. (2010) Orchestrating ribosomal activity from inside: effects of the nascent chain on the peptidyltransferase centre. *Biochem. Soc. Trans.*, **38**, 1576–1580.
8. Atkins,J.F., Wills,N.M., Loughran,G., Wu,C.Y., Parsawar,K., Ryan,M.D., Wang,C.H. and Nelson,C.C. (2007) A case for ''StopGo'': reprogramming translation to augment codon meaning of GGN by promoting unconventional termination (Stop) after addition of glycine and then allowing continued translation (Go). *RNA*, **13**, 803–810.
9. de Felipe,P., Luke,G.A., Hughes,L.E., Gani,D., Halpin,C. and Ryan,M.D. (2006) *E unum pluribus*: multiple proteins from a self-processing polyprotein. *Trends Biotechnol.*, **24**, 68–75.
10. Ryan,M.D., Donnelly,M.L., Lewis,A., Mehotra,A.P., Wilke,J. and Gani,D. (1999) A model for non-stoichiometric, co-translational protein scission in eukaryotic ribosomes. *Bioorg. Chem.*, **27**, 55–79.
11. Hahn,H. and Palmenberg,A.C. (1996) Mutational analysis of the encephalomyocarditis virus primary cleavage. *J. Virol.*, **70**, 6870–6875.
12. Crooks,G.E., Hon,G., Chandonia,J.M. and Brenner,S.E. (2004) WebLogo: a sequence logo generator. *Genome Res.*, **14**, 1188–1190.
13. Sikorski,R.S. and Heiter,P. (1989) A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics*, **122**, 19–27.
14. Schagger,H. and von Jagow,G. (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDA. *Anal. Biochem.*, **166**, 368–379.
15. de Felipe,P., Hughes,L.E., Ryan,M.D. and Brown,J.D. (2003) Co-translational, intraribosomal cleavage of polypeptides by the foot-and-mouth disease virus 2A peptide. *J. Biol. Chem.*, **278**, 11441–11448.
16. Varshavsky,A. (2011) The N-end rule pathway and regulation by proteolysis. *Protein Sci.*, **20**, 1298–1345.
17. Bartel,B., Wunning,I. and Varshavsky,A. (1990) The recognition component of the N-end rule pathway. *EMBO J.*, **9**, 3179–3189.
18. Chaudhuri,B., Ingavale,S. and Bachhawat,A.K. (1997) *apd1+*, a gene required for red pigment formation in *ade6* mutants of *Schizosaccharomyces pombe*, encodes an enzyme required for glutathione biosynthesis: a role for glutathione and a glutathione-conjugate pump. *Genetics*, **145**, 75–83.
19. Donnelly,M.L., Gani,D., Flint,M., Monaghan,S. and Ryan,M.D. (1997) The cleavage activities of aphthovirus and cardiovirus 2A proteins. *J. Gen. Virol.*, **78**, 13–21.
20. Nakatogawa,H. and Ito,K. (2002) The ribosomal exit tunnel functions as a discriminating gate. *Cell*, **108**, 629–636.
21. Vazquez-Laslop,N., Thum,C. and Mankin,A.S. (2008) Molecular mechanism of drug-dependent ribosome stalling. *Mol. Cell*, **30**, 190–202.
22. Spevak,C.C., Ivanov,I.P. and Sachs,M.S. (2010) Sequence requirements for ribosome stalling by the arginine attenuator peptide. *J. Biol. Chem.*, **285**, 40933–40942.
23. Hahn,H. and Palmenberg,A.C. (2001) Deletion mapping of the encephalomyocarditis virus primary cleavage site. *J. Virol.*, **75**, 7215–7218.
24. Seidelt,B., Innis,C.A., Wilson,D.N., Gartmann,M., Armache,J.P., Villa,E., Trabuco,L.G., Becker,T., Mielke,T., Schulten,K. *et al.* (2009) Structural insight into nascent polypeptide chain-mediated translational stalling. *Science*, **326**, 1412–1415.
25. Bhushan,S., Meyer,H., Starosta,A.L., Becker,T., Mielke,T., Berninghausen,O., Sattler,M., Wilson,D.N. and Beckmann,R. (2010) Structural basis for translational stalling by human cytomegalovirus and fungal arginine attenuator peptide. *Mol. Cell*, **40**, 138–146.
26. Bhushan,S., Hoffmann,T., Seidelt,B., Frauenfeld,J., Mielke,T., Berninghausen,O., Wilson,D.N. and Beckmann,R. (2011) SecM-stalled ribosomes adopt an altered geometry at the peptidyl transferase center. *PLoS Biol.*, **9**, e1000581.
27. Brown,J.D. and Ryan,M.D. (2010) Ribosome ''Skipping'': ''Stop-Carry On'' or ''StopGo'' Translation. In: Atkins,J.F. and Gesteland,R.F. (eds), *Recoding*. Springer, New York, pp. 101–121.
28. Ryan,M.D. and Drew,J. (1994) Foot-and-mouth disease virus 2A oligopeptide mediated cleavage of an artificial polyprotein. *EMBO J.*, **13**, 928–933.
29. Barbknecht,M. (2009) Characterisation of a unclassified virus and survey for its presence in Wisconsin Bluegill populations. *MSc Thesis*. University of Wisconsin-La Crosse, La Crosse.