Protein-directed ribosomal frameshifting temporally regulates gene expression

Sawsan Napthine1,*, Roger Ling1,*, Leanne K. Finch1,*, Joshua D. Jones1,*, Susanne Bell1, Ian Brierley1 & Andrew E. Firth1

Programmed −1 ribosomal frameshifting is a mechanism of gene expression, whereby specific signals within messenger RNAs direct a proportion of translating ribosomes to shift −1 nt and continue translating in the new reading frame. Such frameshifting normally occurs at a set ratio and is utilized in the expression of many viral genes and a number of cellular genes. An open question is whether proteins might function as trans-acting switches to turn frameshifting on or off in response to cellular conditions. Here we show that frameshifting in a model RNA virus, encephalomyocarditis virus, is trans-activated by viral protein 2A. As a result, the frameshifting efficiency increases from 0 to 70% (one of the highest known in a mammalian system) over the course of infection, temporally regulating the expression levels of the viral structural and enzymatic proteins.

1 Division of Virology, Department of Pathology, University of Cambridge, Cambridge CB2 1QP, UK. * These authors contributed equally to this work.

Correspondence and requests for materials should be addressed to I.B. (email: ib103@cam.ac.uk) or to A.E.F. (email: aef24@cam.ac.uk).
Programmed −1 ribosomal frameshifting (−1 PRF) is utilized in the expression of many viral genes and a number of cellular genes. It may also play a widespread role in fine-tuning gene expression via the stimulation of nonsense-mediated mRNA decay pathways. Sites of −1 PRF generally comprise a ‘slippery’ sequence (at which the change in reading frame occurs) and a 3′-adjacent stimulatory mRNA structure. In eukaryotes, the slippery sequence fits a consensus heptanucleotide motif X_YYY_ZYY, where XXX is any three identical nucleotides (although certain exceptions occur, such as GUU); YYY represents AAA or UUU; Z represents A, C or U, and underscores separate zero-frame codons. In the tandem slippage model, the ribosomal P-site tRNA anticodon re-pairs from XXY to XXX, and the A-site anticodon re-pairs from YYY to YYY, thus allowing perfect re-pairing except at wobble positions. The efficiency of PRF is influenced by the identity of the slippery site nucleotides but is normally less than 1% in the absence of additional stimulatory elements. Thus, most known instances of eukaryotic −1 PRF are stimulated (typically to a level of 5–45%, depending on the particular case) by the presence of a 3′ stable RNA structure, such as a pseudoknot or stem–loop, separated from the slippery heptanucleotide by a ‘spacer’ region of 5–9 nt. Structures of this type are thought to be located at the mRNA unwinding site of the ribosome entrance channel when their stimulatory effect is exerted. How the stimulatory RNAs function to promote −1 PRF is still uncertain, but accumulating evidence from prokaryotic counterparts indicates that the RNA structure impedes back rotation of the ribosomal small subunit, trapping the ribosome in a rotated or hyper-rotated state. This stalled state can be resolved either via spontaneous unwinding of the structure or via −1 PRF, which, by repositioning the structure within the mRNA entrance channel, allows for more efficient unwinding by the ribosome.

Like other members of the family Picornaviridae, encephalomyocarditis virus (EMCV; genus Cardiovirus) has a single-stranded RNA genome of positive polarity, which also serves as an mRNA. The genome is polyadenylated but lacks a 5′ cap and translation initiation is mediated by an internal ribosome entry site (IRES) within the lengthy 5′ UTR. There is a single long open reading frame whose translation produces a polypeptide that is proteolytically cleaved, mainly by the virus-encoded 3C protease, to produce the structural and enzymatic proteins (Fig. 1a). For more than 50 years, EMCV has been used as a model system for investigating molecular virology, virus–host interactions and eukaryotic protein synthesis.

Studies of EMCV, besides poliovirus and foot-and-mouth disease virus, also led to the discovery of internal ribosome entry and StopGo co-translational separation, with the former now thought to be relevant for some key cellular genes and the latter an important tool in biotechnology. Recently, we uncovered a further unusual aspect of EMCV translation, namely that a previously undetected −1 PRF site in an internal region of the polyprotein ORF directs a proportion of ribosomes into a short overlapping ORF resulting in the production of a ‘trans-frame’ protein, 2B* (Fig. 1a). Our previous work suggested that the PRF mechanism in EMCV, and its relative Theiler’s murine encephalomyelitis virus (TMEV), is atypical due to the apparent absence of an appropriately positioned stimulatory RNA structure, and a failure to reconstitute PRF outside of the context of virus infection.

Here, we describe the discovery of a protein trans-activator of EMCV frameshifting, viral protein 2A. We show that 2A binds to an RNA stem–loop beginning 14 nt downstream of the slippery sequence, forming an RNA:protein complex that induces highly efficient PRF. Through ribosome profiling, we investigate protein synthesis in the natural context of EMCV infection and find, remarkably, that frameshifting is temporally regulated, increasing from negligible levels at early time-points of infection to an efficiency of ~70% at late time-points. Thus, frameshifting serves as a control mechanism to modulate the relative levels of viral structural and non-structural protein synthesis during the replicative cycle. At early time points, ribosome progression to the replicase coding sequences at the 3′ end of the genome is uninterrupted, but, as the replication cycle proceeds, frameshifting progressively diverts ribosomes from the polyprotein ORF to the overlapping 2B* ORF, downregulating replicase translation. These experiments identify and elucidate a mechanism whereby the efficiency of PRF can be temporally regulated, in contrast to the fixed efficiency levels of canonical −1 PRF.

**Results**

Ribosome profiling shows a temporal shift in PRF efficiency. To investigate PRF efficiency in the context of virus infection, we used ribosome profiling (Ribo-Seq) to directly monitor the density of ribosomes upstream and downstream of the frameshift site. Ribosome profiling is a recently developed technique that enables global footprinting of translating ribosomes in cells via high-throughput sequencing of ribosome-protected fragments (RPFs) (Fig. 1b). We infected murine L929 cells with wild-type (WT) or shift site mutant (SS) viruses and harvested cells at 2, 4, 6 and 8 h post infection (p.i.). In the SS mutant, the WT shift site G_GUU_UUU is mutated to A_GUG_UUU (Fig. 1c) to inhibit PRF. Ribo-Seq libraries were prepared from each sample, deep sequenced, and the resulting reads mapped to host and viral genomes. Ribo-Seq quality was assessed as described previously (Supplementary Fig. 1).

Figure 1d shows the distributions of RPFs on the WT and SS virus genomes at 4 and 8 h p.i. (see Supplementary Fig. 2a for 2 and 6 h p.i.). For WT virus, a substantial decrease in mean ribosome density was observed to occur after the 2B* ORF at 6 and 8 h p.i., indicating that the majority of ribosomes translating the viral genome frameshift into the 2B* ORF and terminate at the 2B* stop codon. For the SS mutant, however, there is no significant drop in RPF density at the 2B* stop codon, consistent with inhibited frameshifting in this mutant. PRF efficiencies were estimated from the WT ratio of downstream to upstream RPF densities, normalized by the SS mutant to control for variation in translational speed, technical biases and ribosome pausing between different genomic locations. PRF efficiencies were not obviously different from zero at 2 and 4 h p.i., but at 6 and 8 h p.i. were remarkably high, at 64% and 69%, respectively (Fig. 1e and Supplementary Fig. 2b). Consistent with the absence of measurable PRF at 2 and 4 h p.i., a noticeable ribosomal pause that was observed in the 2B* region at 6 and 8 h p.i. for WT virus, is absent at 2 and 4 h p.i. and in the SS mutant (Fig. 1d and Supplementary Fig. 2a). The effect of this pause is uncertain though, in common with many other ribosomal pauses, it may be nascent peptide mediated. Extraordinarily, from 6 h p.i. RPFs accumulated to high levels on the mutated shift site (Fig. 1d and Supplementary Fig. 2a). In contrast, only a modest accumulation was apparent on the WT shift site. The position of RPFs in this region indicate that the pause initiates when the UUU of the shift site is in the ribosomal A-site and, in these samples, continues for ~3 codons thereafter (Supplementary Fig. 2c,d), although such broadening of the pause is likely an artefact resulting from ribosome run-on during sample preparation. The total excess accumulation of RPFs over these four codons at 6 and 8 h p.i. relative to 4 h p.i. is ~5-fold greater in SS than in WT.

Ribosome profiling of eukaryotic systems typically has the characteristic that mappings of the 5′ end positions of RPFs to coding sequences reflect the triplet periodicity (herein referred to...
as ‘phasing’) of genetic decoding. Figure 1f and Supplementary Fig. 2 show the polyprotein-frame codon positions to which the 5′ ends of RPFs map for regions upstream of 2B*, within 2B*, and downstream of 2B*. For the upstream and downstream regions, RPF 5′ ends map mainly to the first and third positions of polyprotein-frame codons. In contrast, within the 2B* region, RPF 5′ ends for WT virus at 6 and 8 h.p.i. map mainly to the third and second positions of polyprotein-frame codons, consistent with a mixture of ribosomes translating the two overlapping reading frames in this region. For the SS mutant, there is no change in the phasing of RPF 5′ end positions in the overlap region (Fig. 1f), consistent with inhibited PRF in this mutant.

**An RNA stem–loop is essential for PRF and ribosome pausing.**

Previously, a downstream stem–loop structure, separated from the frameshift site by a 13-nt ‘spacer’, was identified bioinformatically. This positioning is inconsistent with canonical mRNA structure stimulators of −1 PRF, which are separated from the shift site by just 5−9 nt. The predicted stem–loop is conserved in different isolates of EMCV and in TMEV, with compensatory substitutions (i.e. paired substitutions that preserve the predicted base-pairings) further supporting its biological relevance (Supplementary Fig. 3a). RNA structure probing also supported the stem–loop (Supplementary Fig. 3b,c). To assess a role for the stem–loop in PRF and/or ribosome pausing, we generated a stem–loop mutant virus, WT-SL, and a corresponding shift site mutant, SS-SL, and performed ribosome profiling at 8 h.p.i. (Fig. 2a,b)23. For LV-WT, RPF phasing in the overlap region was the frameshift signal is located just downstream of the polyprotein ORF (Figs 1a and 2a). Separation between 2A- and 2B-encoding regions of the EMCV frameshift signal is located just downstream of the polyprotein ORF (Figs 1a and 2a). Separation between 2A- and 2B-encoding regions of the EMCV frameshift signal is located just downstream of the polyprotein ORF (Figs 1a and 2a). Separation between 2A- and 2B-encoding regions of the EMCV frameshift signal is located just downstream of the polyprotein ORF (Figs 1a and 2a). Separation between 2A- and 2B-encoding regions of the EMCV frameshift signal is located just downstream of the polyprotein ORF (Figs 1a and 2a). Separation between 2A- and 2B-encoding regions of the EMCV frameshift signal is located just downstream of the polyprotein ORF (Figs 1a and 2a). Separation between 2A- and 2B-encoding regions of the EMCV frameshift signal is located just downstream of the polyprotein ORF (Figs 1a and 2a). Separation between 2A- and 2B-encoding regions of the EMCV frameshift signal is located just downstream of the polyprotein ORF (Figs 1a and 2a). Separation between 2A- and 2B-encoding regions of the EMCV frameshift signal is located just downstream of the polyprotein ORF (Figs 1a and 2a). Separation between 2A- and 2B-encoding regions of the EMCV frameshift signal is located just downstream of the polyprotein ORF (Figs 1a and 2a). Separation between 2A- and 2B-encoding regions of the EMCV frameshift signal is located just downstream of the polyprotein ORF (Figs 1a and 2a). Separation between 2A- and 2B-encoding regions of the EMCV frameshift signal is located just downstream of the polyprotein ORF (Figs 1a and 2a). Separation between 2A- and 2B-encoding regions of the EMCV frameshift signal is located just downstream of the polyprotein ORF (Figs 1a and 2a). Separation between 2A- and 2B-encoding regions of the EMCV frameshift signal is located just downstream of the polyprotein ORF (Figs 1a and 2a). Separation between 2A- and 2B-encoding regions of the EMCV frameshift signal is located just downstream of the polyprotein ORF (Figs 1a and 2a). Separation between 2A- and 2B-encoding regions of the EMCV frameshift signal is located just downstream of the polyprotein ORF (Figs 1a and 2a). Separation between 2A- and 2B-encoding regions of the EMCV frameshift signal is located just downstream of the polyprotein ORF (Figs 1a and 2a). Separation between 2A- and 2B-encoding regions of the EMCV frameshift signal is located just downstream of the polyprotein ORF (Figs 1a and 2a). Separation between 2A- and 2B-encoding regions of the EMCV frameshift signal is located just downstream of the polyprotein ORF (Figs 1a and 2a). Separation between 2A- and 2B-encoding regions of the EMCV frameshift signal is located just downstream of the polyprotein ORF (Figs 1a and 2a). Separation between 2A- and 2B-encoding regions of the EMCV frameshift signal is located just downstream of the polyprotein ORF (Figs 1a and 2a). Separation between 2A- and 2B-encoding regions of the EMCV frameshift signal is located just downstream of the polyprotein ORF (Figs 1a and 2a). Separation between 2A- and 2B-encoding regions of the EMCV frameshift signal is located just downstream of the polyprotein ORF (Figs 1a and 2a). Separation between 2A- and 2B-encoding regions of the EMCV frameshift signal is located just downstream of the polyprotein ORF (Figs 1a and 2a). Separation between 2A- and 2B-encoding regions of the EMCV frameshift signal is located just downstream of the polyprotein ORF (Figs 1a and 2a). Separation between 2A- and 2B-encoding regions of the EMCV frameshift signal is located just downstream of the polyprotein ORF (Figs 1a and 2a).
not greatly different from WT (Fig. 2d), and the calculated PRF efficiency (62%) was only slightly lower than WT (70%) (Fig. 2e). The modest difference may be due to slower replication in the mutant leading to a lag in achieving maximum PRF efficiency, or may indicate a more direct (albeit modest) stimulatory effect of StopGo on PRF, likely related to proper processing of the C terminus of 2A (see below). On the other hand, for the mutant LV-SS-SL these features were similar to the SS-SL mutant. Thus we found no evidence for ribosome drop-off induced by StopGo alone.

**Virus protein 2A binds the RNA stem–loop.** Having observed that PRF in EMCV increases from 0 to 70% over the course of infection and depends critically on an RNA stem–loop structure positioned unusually far 3′ of the frameshift site, we reasoned that some virus-induced trans-acting protein might interact with the stem–loop to promote PRF. To test this, we used the RiboTrap system in which RNA transcripts are labelled with 5-bromo-uridine to allow immunopurification of bound complexes. A short RNA transcript containing the EMCV stem–loop was found to bind an ~16 kDa protein from StopGo mutation (LV) Mutated shift site (SS) Stem–loop mutations (SL) 

![Diagram](https://example.com/diagram.png)

**Figure 2 | An RNA stem–loop is required for frameshifting.** (a) Mutations introduced to prevent PRF (SS) or StopGo-mediated co-translational separation at the C-terminus of 2A (LV) or to destabilize the predicted RNA stem–loop (SL). (b) Ribo-Seq RPF densities in reads per million mapped reads (RPM) on WT and mutant virus genomes at 8 h p.i. WT and SS samples are biological repeats of the 8 h p.i. samples in Fig. 1. (c) Ratio of SS and SS-SL RPF distributions after smoothing with a 31-nt running-mean filter. (d) Phasing of RPFs mapping upstream of 2B*, within the 2B/2B* overlap region, and downstream of 2B*. (e) Ratio of downstream to upstream RPF densities in the different viruses divided by the corresponding ratio in SS mutant virus.
EMCV-infected cell lysates that was not observed with mock-infected cell lysates or with a control RNA containing a scrambled version of the EMCV stem–loop sequence (Supplementary Fig. 6a). Mass spectrometric analysis of this product revealed it to be the viral protein 2A (Fig. 1a and Supplementary Fig. 6b).

To confirm this interaction, we expressed and purified recombinant 2A protein and performed electrophoretic mobility shift assays (EMSAs) with a 64-nt 32P-labelled RNA containing the EMCV frameshift site and stem–loop. Two main RNA:protein complexes were observed, with the more slowly migrating species accumulating as the amount of 2A was increased (Fig. 3a), confirming that 2A does indeed bind the EMCV PRF signal. We estimated the $K_d$ of the interaction to be $\sim 3–4.5 \mu M$ (Fig. 3b). A CCC triplet in the RNA loop is highly conserved between different isolates of EMCV and TMEV (Supplementary Fig. 3a), suggesting a potential role in 2A binding. To investigate this, we mutated the middle C to U (C46U; Fig. 3c) and tested whether cold WT RNA or cold mutant RNA could compete with the 32P-labelled WT RNA for 2A binding. Whereas the WT RNA was able to compete well, greatly diminishing radiolabelled RNA:protein complexes at increasing molar excess, the C46U mutant was unable to compete, indicating that this single-nucleotide mutation had prevented 2A binding (Fig. 3d). Similarly, RNAs with stem-disrupting mutations (G52U and SL3; Fig. 3c) were also unable to compete with WT RNA (Fig. 3e), indicating that the presence of at least part of the stem–loop duplex is also important for 2A binding (and not just for proper positioning of the RNA:protein complex for PRF stimulation). A direct interaction between 2A and RNA would likely involve a cluster of positively charged residues on 2A. By inspecting an alignment of cardiovirus 2A sequences we identified a conserved linear basic cluster (Supplementary Fig. 6c). We generated a recombinant mutant 2A (hereafter 2A-mut) by changing R95 and R97 to alanines (Supplementary Fig. 6d,e). EMSA analysis confirmed that 2A-mut was unable to bind to the EMCV PRF signal (Fig. 3f). Further, when introduced into the virus genome, the
from Fsp immunoblotting using antibodies to 2A (green) and tubulin (magenta; loading control).

Virus protein 2A trans-activates ribosomal frameshifting. Next we asked whether the 2A protein is sufficient to stimulate PRF in the absence of viral infection. To test this we titrated recombinant 2A into a wheat germ (WG) in vitro translation system programmed with a reporter mRNA containing the EMCV PRF signal (Fig. 5a). Increasing amounts of 2A led to a modest general inhibition of cap-dependent translation, as reported previously24,25. When the reporter was translated in the absence of 2A, only the non-frameshift product was observed (Fig. 5b, lane DB). However, in the presence of 2A, efficient PRF was observed, to a level of ~20% with increasing amounts of 2A (Fig. 5b,c). In control translations, PRF was abolished upon mutating the frameshift site (SS), or the stem–loop and CCC triplet (SL), or upon replacing 2A with 2A-mut (Fig. 5d).

R95 and R97 substitutions resulted in a protein expression pattern similar to that of SS mutant virus, indicating that the 2A mutations had inhibited PRF (Fig. 4a–c). Consistent with the increase in frameshifting efficiency, the amount of 2A in cells infected with WT virus increases dramatically between 4 and 6 h p.i. (Fig. 4d).

Figure 4 | Mutating 2A knocks out frameshifting. (a) Plaque morphology of WT, SS and 2A-mut viruses on BHK-21 cells (see also ref. 25). (b) Metabolic labelling of BHK-21 cells mock-infected or infected with WT, SS or 2A-mut viruses. Positions of EMCV proteins are indicated. (c) Mean ratio of measureable polyprotein products encoded downstream of the frameshift site to products encoded upstream of the frameshift site, normalized by the SS mutant. Bars show means ± s.d. of three biological replicates (green crosses). (d) Time course of 2A expression during WT virus infection assessed by immunoblotting using antibodies to 2A (green) and tubulin (magenta; loading control).

Figure 5 | Analysis of the EMCV frameshift signal. (a) Mutations introduced into the EMCV PRF signal in reporter plasmid pDluc. (b–e) RNAs derived from FspI-cut plasmid were translated in WG extract in the presence of (b) increasing concentrations of 2A or (d,e) 3 μM 2A or 2A-mut, or 2A dialysis buffer (DB). Products generated by ribosomes that do not frameshift (stop) or that enter the –1 reading frame (–1 FS) are indicated. M and IFC indicate markers and the in-frame control, respectively. Panel e shows the PRF efficiencies for b; the dashed line indicates the 2A concentration used in (d,e).
Changing the loop-region CCC triplet to UUU (3CU), CUC (C46U) or CAC (C46A) also inhibited PRF (Fig. 5e). The SL5‘ and SL3‘ mutations of Figs 3c and 5a substantially reduced PRF (to 4% and 1%, respectively) whereas the SL5‘3‘ mutations (designed to restore the stem–loop structure but with altered base pairings) restored PRF to WT levels (Fig. 5e). Further, increasing (+ U) or decreasing (− U) the spacer length by 1 nt reduced PRF to 6 and 1%, respectively (Fig. 5e). Similar results were observed in the rabbit reticulocyte lysate (RRL) in vitro translation system (Supplementary Fig. 7).

The in vitro translation systems were also used to further investigate ribosome pausing at the frameshift site (Fig. 6 and Supplementary Fig. 8). The extent of pausing was assessed by comparing the levels of a translational intermediate corresponding to pausing at the EMCV shift site with that of the full-length polypeptide produced during a time course in which translation was synchronized by the addition of edeine, a potent inhibitor of ribosome pausing. When the spacer length was increased by 1 nt (SS + U and WT + U) pausing was observed in the presence of 2A but was much less pronounced than for SS RNA (Supplementary Fig. 8e,f). The appearance of the pause product is transitory (albeit spread over several minutes for WT and many minutes for SS; Fig. 6b,c), consistent with its identity as a genuine intermediate rather than a dead-end product. There was no

Figure 6 | Ribosomal pausing at the EMCV frameshift signal. (a) Mutations introduced into the EMCV PRF signal in reporter plasmid pPS0. (b–e) RNAs derived from Avall-cut plasmids were translated in WG extract, after 5 min further initiation was halted by the addition of edeine, and aliquots were removed at various times and analysed by SDS–PAGE. Lanes M and C show markers and the expected size of the ribosomal pause product, respectively. Translations were supplemented with 1 μM 2A (b–d) or 2A-mut (e). As well as the full-length product and the transient pausing product, a frameshift product is produced for WT RNA only (e).

Figure 7 | Comparison of −1 frameshifting stimulators. (a) Nearly all known cases of −1 PRF are stimulated by an mRNA structure, such as a pseudoknot or stem–loop, separated from the slippery heptanucleotide shift site by a spacer sequence of 5–9 nt, leading to a constant ratio of frameshift to non-frameshift product. (b) In contrast, frameshifting in EMCV is stimulated by a protein:RNA complex positioned at the leading edge of the ribosome when the decoding centre is on the shift site. Increasing levels of viral protein 2A result in a 0 to 70% switch in frameshifting efficiency between early and late timepoints of infection.
evidence for significant ribosomal drop-off occurring at the WT shift site (Fig. 6c and Supplementary Fig. 8b, 60-min lane) though we could not rule out a small amount of drop-off occurring when the shift site was mutated due to the long half-life of the pause product for these constructs (Fig. 6b and Supplementary Fig. 8a).

Discussion

We have shown that PRF in EMCV is extraordinarily efficient, and, in stark contrast to other cases of 1 PRF such as from HIV or SARS coronavirus, varies dramatically over the course of infection from 0 to ~70% and depends crucially on a trans-activating virus protein (Fig. 7). Early in infection the efficiency of PRF is negligible, allowing efficient translation of the 3'-encoded replication proteins. However, late in infection and presumably as a result of increasing concentrations of cytoplasmic 2A, PRF becomes highly efficient, resulting in >3-fold downregulation of replication protein synthesis. This provides an elegant and economic solution to facilitate maximal build up of replication capacity early in infection, while providing a translational bias that favours virion production at late timepoints, and raises the question of whether other 'single-polypeptide' viruses such as poliovirus and hepatitis C might also use non-canonical translational mechanisms to temporally regulate replication protein synthesis. These results provide a satisfying conclusion to observations by Paucha et al. who, 40 years ago, observed approximately twice as much capsid as noncapsid viral protein in EMCV-infected cells, specifically at late timepoints. Although they could not then provide a mechanism, these authors showed extraordinary insight—long before the EMCV genome was sequenced—by proposing a model in which, to quote, 'some viral-specific protein acts as a specific termination factor capable of causing premature termination of translation at a site located near the midpoint of the viral RNA'. We can now confirm that this effect is mediated at the translational level, and depends on protein-stimulated PRF rather than ribosomal drop-off at StopGo or at other sites during polyprotein synthesis.

Together with 2 PRF for nsp2TF expression in members of the family Arteriviridae, where we recently demonstrated that PRF is critically dependent on the arteriviral protein nsp1β and host poly(C) binding proteins interacting with a 3'-C-rich motif separated from the frameshift site by a 10-nt spacer, this is one of only two known cases of protein-stimulated PRF. As 2A and nsp1β are viral proteins, they cannot be utilized for normal cellular gene expression. Nonetheless there may be other cases of protein-stimulated PRF and the two extant examples suggest that the protein binding sites can be positioned just outside the mRNA entrance channel at the onset of frameshifting. Cellular analogues would have been missed by previous bioinformatic searches for PRF sites that have relied on the presence of a predicted RNA structure beginning closer to a slippery heptanucleotide shift site.

An association between ribosomal pausing and RNA-structure-stimulated 1 PRF has been previously reported though it has been unclear to what extent pausing is required for 1 PRF as opposed to being an unavoidable side effect. Of note, several previous pausing analyses used reporters with a mutated shift site. Similar to recent analyses of RNA structure-stimulated 1 PRF in a cell-free Escherichia coli system, we observed much more pronounced pausing when the shift site was mutated than for an intact shift site (Supplementary Fig. 2c,d in cells and Fig. 6b,c in vitro), suggesting that, by making a 1 nt shift, ribosomes are able to more efficiently remove 2A and unwind the stem–loop than if they remain in the original frame. Similar to that study, but in contrast to recent single-molecule FRET work, the EMCV pause starts when the P- and A-sites are on the shift site, not upstream. The reason for the significantly lower PRF efficiency observed in vitro (maximum ~20%) compared to virus infection (maximum ~70%) remains unknown, but possibilities include incomplete activity of the recombinant 2A, effects of more distal RNA sequences, or differences in translational environment, such as salt, temperature, pH and ribosome loading. In the absence of artificially processed 2A, PRF in the StopGo (LV) mutants would be stimulated by 2A–2B (or cleaved versions thereof) and/or 2A–2B+ and this may explain the slightly reduced PRF level in LV-WT (Fig. 2e). The stem–loop:2A complex likely acts as an analogue of an RNA structure stimulator, except that it is regulatable. At this stage it is not known if it is simply an obstacle (e.g. to ribosome subunit rotation) or has some more specific interaction with the ribosome.

This report describes an example of a temporally regulated PRF switch. A second example of viral protein-stimulated PRF has been noted recently in arenaviruses, and the RNA-stimulated PRF signal has been discovered in the mRNA encoding the HIV-1 co-receptor CCR5 (ref. 4). The recent identification of these unrelated examples of trans-activated PRF signals indicates such signals may be more widespread in viruses and their hosts, and that regulated expression through the stimulation of 1 PRF will become an established paradigm of gene expression.

Methods

Cells, recombinant viruses and plasmids. Cell lines were obtained from the European Collection of Authenticated Cell Cultures (ECACC) and tested for mycoplasma by PCR (e-Mycoplus Mycoplasma PCR Detection Kit; iNtRON Biotechnology). In addition, the sequenced libraries (L929 cells) were queried for mycoplasma sequences. WT and mutant viruses are based on the EMCV subtype mengo virus cDNA, pMC10, developed by Ann Palminenberg (University of Wisconsin-Madison)32. The parental (WT) sequence used is similar to GenBank accession DQ294633.1, but a 55-nt poly(C) tract in the 5' UTR is deleted and there are 13 single-nucleotide differences (A2669C, G3044C, C3371T, A4910C, G4991A, G5156T, G5289A, G5314C, G5315A, A5844C, G6266A, G6990A, A6992G, DQ294633.1 coordinates). Excepting the preceding sentence, nucleotide coordinates are given with respect to vMC1. WT, SS and LV-WT viruses were a kind gift from Gary Loughran (University College Cork)7,25. All constructs were prepared by standard PCR mutagenesis and recombinant DNA techniques and subcloned regions altered by mutagenesis were verified by DNA sequencing. All viruses were able to replicate in cell culture. The SS and SL mutations do not alter the polypeptide amino acid sequence.

In vitro frameshifting at the EMCV PRF signal. For in vitro frameshifting assays, we cloned a 105-nt sequence containing the G_GUU_UUU shift site flanked by 12 nt upstream and 86 nt downstream, or mutant derivatives, into the dual luciferase plasmid pDLuc at the XhoI/BglII sites33. The sequence was inserted between the Renilla and firefly luciferase genes so that firefly luciferase expression is dependent on 1 PRF. For ribosomal pausing analysis, the EMCV shift site flanked by 5 nt upstream and 93 nt downstream, or mutant derivatives, were cloned into pPS0 at the XhoI/PsiI sites27. For the expression of recombinant 2A in E. coli, the 2A coding sequence was amplified from pMC10 and cloned into pGEX-6P-2 (GE Healthcare) at the BamHI/XhoI sites. The expressed 2A, following removal of the glutathione-S-transferase moiety by PreScission Protease (a kind gift from Stephen Graham, University of Cambridge), has an additional 4 and 13 vector-derived residues at its N- and C-termini, respectively.

In vitro transcription and generation of recombinant virus. RNA was transcribed using the Megascript T7 kit (Ambion) from BsmHI-linearized plasmids. Reactions were phenol/chloroform extracted, the RNA desalted by centrifugation through a NucAway Spin Column (Ambion) and concentrated by ethanol precipitation. Purified RNAs were used to transfect 35-mm dishes of I929 or BHK-21 cells using 1.2 μg RNA and 4 μl DMRIE-C reagent according to the manufacturer’s instructions. Transfected cells were cultured in Dulbecco’s modified Eagle’s medium (DMEM) with high glucose and 1% fetal bovine serum (FBS) for 1–5 days depending upon how rapidly cytopathic effect developed. Cultures were subjected to three rounds of freeze–thawing, cell debris removed by centrifugation for 5 min at 4,000g and the supernatant stored in aliquots at −80 °C.
Plaque assays. BHK-21 cells at 90% confluence in six-well plates were infected with serial dilutions of virus stocks. Cells were washed with serum-free medium, overlaid with virus inoculum and incubated at 37 °C. Infected and uninfected wells were removed and replaced with 1.5% low melting point agarose (Invitrogen) containing DMEM plus 2% FBS. After 40 h incubation at 37 °C, cells were fixed with formal saline and stained with 0.1% toluidine blue.

Metabolic labelling and calculation of PRF efficiencies. BHK-21 cells were infected at a multiplicity of infection (MOI) of ~10 in a volume of 150 μl in 24-well plates. After 1 h the inoculum was replaced with 1 ml DMEM containing 1% FBS. At 7 h p.i., cells were incubated for 1 h in methionine- and serum-free DMEM, and radiolabelled from 8 to 9 h p.i. with [35S]methionine at 100 μCi/ml (~1,100 GBq mmol⁻¹) in methionine-free medium. Cells were scraped into the medium, pelleted at 13,000g for 1 min, washed twice by resuspension in 1 ml of ice-cold phosphate-buffered saline (PBS) and pelleted for 2 min at 13,000g. Cell pellets were lysed in 35 μl x SDS-PAGE sample buffer and boiled for 5 min before analysis by SDS-PAGE. Dried gels were exposed to X-ray films or to phosphorimager storage screens. Image analysis was carried out using ImageQuant TL 7.0, and the radioactivity in virus-specific products quantified.

The intensity for each WT virus product was measured, normalized by methionine content and then by the mean value for VP0, VP3 and VP1 to control for lane loading. Next, to factor out differences in protein turnover besides unquantified processing intermediates, for each biological replicate the WT values for VP0, VP3, VP1, 2C, 3A+B, 3C+3D and 3D+3CD were normalized by corresponding values for SS mutant virus. Then the normalized values for 2C, 3A+B, 3C+3D and 3D+3CD (i.e. products encoded downstream of the frameshift site) and divided by the average of the values for VP0, VP3 and VP1 (i.e. products encoded upstream of the frameshift site). This gives an estimate of the fraction of ribosomes that avoid a ~1 PRF (Fig. 4c). One minus this value estimates the PRF efficiency.

Ribosome profiling library preparation. L929 cells in 60-mm dishes were infected at an MOI of ~10 in an initial volume of 1.5 ml DMEM with 1% FBS. After 1 h adsorption at 37 °C an additional 3.5 ml was added and incubation continued. At the appropriate time points, cells were treated with cycloheximide (Sigma-Aldrich; 10 μg/ml; 2 h). Cells were rinsed with 5 ml of ice-cold PBS, the dishes submerged in a reservoir of liquid nitrogen for 10 s transferred to dry ice and 400 μl of lysis buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM MgCl₂, 1 mM DTT, 1% Triton X-100, 100 μg ml⁻¹ cycloheximide and 25 U ml⁻¹ TURBO DNase (Life Technologies)) dripped onto the cells. The cells were scraped extensively to ensure lysis, collected and triturated with a 26-G needle 10 times. Lysates were clarified by centrifugation for 20 min at 13,000g at 4 °C, the supernatants recovered and stored in liquid nitrogen. Cell lysates were subjected to Ribo-Seq and RNA-Seq. The methodologies employed were based on the original protocols of Ingolia et al.34,35 except library amplification were constructed using a small RNA sequencing kit35 adapted to Illumina smallRNA v2 to allow multiplexing. Ribosome-based RNA-Seq libraries were used for RNA-Seq libraries (Illumina, Santa Clara, CA). RNA-seq libraries were sequenced on the Illumina HiSeq 2000 platform at the Beijing Genomics Institute (round 1) and the Illumina NextSeq platform at the Department of Biochemistry, University of Cambridge (round 2). Sequencing data have been deposited in ArrayExpress (http://www.ebi.ac.uk/arrayexpress) under the accession number E-MTAB-5206.

Structure probing. Short, 33P-labelled RNAs (111 nt) containing the EMCV PRF region (shift site plus 32 nt upstream and 72 nt downstream) were prepared by T7 transcription of a PCR product generated using primers flanking the PRF region, indicating the 5′ primer containing the 17-ribosome precursor sequence. The Transcripts (5 μg) were dephosphorylated with Antarctic phosphatase and then used in a reaction containing 20 μl [γ-33P]ATP (100 μCi ml⁻¹ stock; PerkinElmer), 50 mM Tris pH 7.5, 10 mM MgCl₂, and 5 mM DTT. The labelled RNA was then subjected to a 7% denaturing gel, full-length RNA eluted from the gel slice in 0.5 M NaH₂OAc, 10 mM MgOAc, 1 mM ethylenediaminetetraacetic acid (EDTA) and 0.1% SDS and concentrated by ethanol precipitation. RNA structure probing reactions were performed in a final volume of 50 μl containing ~20,000 c.p.m. 33P [γ-33P] end-labelled transcript, 2 mM MgCl₂, 10 μg pig liver RNA and the relevant enzymatic or chemical probes. Products were analysed on a 10% acrylamide/7 M urea gel.

RiboTrap and mass spectrometry. RNA-binding proteins that may associate with the EMCV PRF signal were screened using the RiboTrap method using a commercial kit (RiboCluster Profiler RiboTrap Kit; Medical & Biological Research, Inc.) under manufacturer’s recommendation (New England BioLabs) and 5 μl of RiboTrap.
Laboratories, Japan)11. WT and scrambled EMCV RNA baits were generated by in vitro transcription from plasmid pBKS/EMCV, which contains 126 nt of EMCV sequence, beginning immediately 3′-adjacent to the GGUUU UUUU shift site. The 145 nt T7 transcript generated from this vector comprises 126 nt of virus translation until free and bound RNA species were resolved, then fixed for 15 min in ice-cold PBS before promptly loading the mix onto 4% acrylamide non-denaturing gels for 3 h at 30°C.

Ribosome pausing assays. WG and RRL in vitro translation reactions (30 μl) were supplemented with 1 μM of 2A, 2A-mut or dialysis buffer, and programmed with mRNAs derived from Assv1-cut pBKS-EMCV-WT or mutant derivatives. Reactions were incubated at 26°C for 3 min prior to the addition of edeine to 5 μM final concentration. Aliquots (1.5 μl) were subsequently withdrawn at set intervals post-edeine addition, mixed with an equal volume of 100 μg ml⁻¹ RNase A in 10 mM EDTA and plated on ice. At the end of the time-course, products were resolved by 12% SDS–PAGE. The expected size of the ribosomal pause product was marked by translating a control mRNA produced from XhoI-cleave pSP80.

Data availability. Sequencing data have been deposited in ArrayExpress (http://www.ebi.ac.uk/arrayexpress) under the accession number E-MTAB-5206. The data that support the findings of this study are available from the corresponding author upon request.

References

1. Dinman, J. D. Mechanisms and implications of programmed translational frameshifting. Wiley Interdiscip. Rev. RNA 3, 661–673 (2012).
2. Caliskan, N., Peske, F. & Rodnina, M. V. Changed in translation: mRNA recoding by -1 programmed ribosomal frameshifting. Trends Biochem. Sci. 40, 265–274 (2015).
3. Atkins, J. F., Loughran, G., Bhatt, P. R., Firth, A. E. & Baranov, P. V. Ribosomal frameshifting and transcriptional slippage: from genetic snaegonography & crytoptography to adventitious use. Nucleic Acids Res. 44, 7007–7018 (2016).
4. Beleu, A. T. et al. Ribosomal frameshifting in the CCR5 mRNA is regulated by mRNAs and the NMD pathway. Nature 512, 265–269 (2014).
5. Brierley, J., Jenner, A. J. & Ingles, S. C. Mutational analysis of the “slippery-sequence” component of a coronavirus ribosomal frameshifting signal. J. Mol. Biol. 227, 463–479 (1992).
6.acks, T., Madhani, H. D., Masiarz, F. R. & Varmas, H. E. Signals for ribosomal frameshifting in the Rea souma virus gag-pol gene. Cell 55, 447–458 (1988).
7. Qu, X. et al. The ribosome uses two active mechanisms to unwind messenger RNA during translation. Nature 475, 118–121 (2011).
8. Chen, J. et al. Dynamic pathways of -1 translational frameshifting. Nature 512, 328–332 (2014).
9. Caliskan, N., Katunin, V. I., Belardinelli, R., Peske, F. & Rodnina, M. V. Programmed -1 frameshifting by kinetic partitioning during impeded translacation. Cell 157, 1619–1631 (2014).
10. Lark, S. K. et al. A segment of the 5′ nontranslated region of ecephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. J. Virol. 62, 2636–2643 (1988).
11. Baltimore, D. & Franklin, R. M. A new ribonucleic acid polymerase appearing after mengovirus infection of L-cells. J. Biol. Chem. 238, 3395–3400 (1963).
12. Baltimore, D., Franklin, R. M. & Callender, J. Mengovirus-induced inhibition of host ribonucleic acid and protein synthesis. Biochim. Biophys. Acta 76, 425–430 (1963).
13. Kegg, I. M., Cohen, N. & Work, T. S. Factors controlling amino acid incorporation by ribosomes from Krebs II mouse ascites-tumour cells. Biochem. J. 98, 826–835 (1966).
14. Pelletier, J. & Sonenberg, N. Internal initiation of translation of eukaryotic mRNAs directed by a sequence derived from poliovirus RNA. Nature 334, 320–325 (1988).
15. Ryan, M. D. & Drew, J. Foot-and-mouth disease virus 2A oligopeptide mediated cleavage of an artificial polyprotein. EMBO J. 13, 928–933 (1994).
16. Hahn, H. & Palmenberg, A. C. Mutational analysis of the ecephalomyocarditis virus primary cleavage. J. Virol. 70, 6870–6875 (1996).
17. Loughran, G., Firth, A. E. & Atkins, J. F. Ribosomal frameshifting into an overlapping gene in the 28 encoding region of the cardiovascular genome. Proc. Natl Acad. Sci. USA 108, E111–E1119 (2011).
18. Finch, L. K. et al. Characterization of ribosomal frameshifting in Thielers’ disease encephalomyovirus. J. Virol. 89, 8580–8589 (2015).
19. Ingolia, N. T., Ghaemmaghami, S., Newman, J. R. & Weissman, J. S. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science 324, 218–223 (2009).

NATURE COMMUNICATIONS | DOI: 10.1038/ncomms15582 | www.nature.com/naturecommunications
20. Ingolia, N. T., Lareau, L. F. & Weissman, J. S. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian poly(A) site selection. *J. Gen. Virol.* 94, 348–353 (2013).

21. Irigoyen, N. et al. High-resolution analysis of coronavirus gene expression by RNA sequencing and ribosome profiling. *PLoS Pathog.* 12, e1005473 (2016).

22. Hussmann, J. A., Patchett, S., Johnson, A., Sawyer, S. & Press, W. H. Understanding biases in ribosome profiling experiments reveals signatures of translation dynamics in yeast. *PLoS Genet.* 11, e1005732 (2015).

23. Loughran, G. et al. Theiler’s murine encephalomyelitis virus contrasts with encephalomyocarditis and foot-and-mouth disease viruses in its functional utilization of the StopGo non-standard translation mechanism. *J. Gen. Virol.* 94, 457–468 (2013).

24. Aminov, A. G., Amineva, S. P. & Palmenberg, A. C. Encephalomyocarditis viral protein 2A localizes to nucleoli and inhibits cap-dependent mRNA translation. *Virus Res.* 95, 45–57 (2003).

25. Groppo, R., Brown, B. A. & Palmenberg, A. C. Mutational analysis of the EMCV 2A protein identifies a novel localization signal and an eIF4E binding site. *Virology* 410, 257–267 (2011).

26. Paucha, E. & Colter, J. S. Evidence for control of translation of the viral genome during replication of Mengo virus and poliovirus. *Virology* 67, 300–305 (1975).

27. Naphine, S. et al. A novel role for poly(C) binding proteins in programmed ribosomal frameshifting. *Nucleic Acids Res.* 44, 5491–5503 (2016).

28. Belew, A. T., Hepler, N. L., Jacobs, J. L. & Dinman, J. D. PRFdb: a database of computationally predicted eukaryotic programmed -1 ribosomal frameshift signals. *BMC Genomics* 9, 339 (2008).

29. Somoyi, P., J., J. E. & Palmenberg, A. C. Ribosomal pausing during translation of an RNA pseudoknot. *Mol. Cell Biol.* 13, 6931–6940 (1993).

30. Lopinski, J. D., Dinman, J. D. & Bruenn, J. A. Kinetics of ribosomal pausing during programmed -1 translational frameshifting. *Mol. Cell Biol.* 20, 1095–1103 (2000).

31. Kontos, H., Naphine, S. & Brierley, I. Ribosomal pausing at a frameshifter RNA pseudoknot is sensitive to reading phase but shows little correlation with frameshift efficiency. *Mol. Cell Biol.* 21, 8657–8670 (2001).

32. Duke, G. M., Osoiro, J. E. & Palmenberg, A. C. Attenuation of Mengo virus through genetic engineering of the 5′ noncoding poly(C) tract. *Nature* 343, 474–476 (1990).

33. Fixsen, S. M. & Howard, M. T. Processive selenocysteine incorporation during synthesis of eukaryotic selenoproteins. *J. Mol. Biol.* 399, 385–396 (2010).

34. Ingolia, N. T., Brag, R. A., Rouskin, S., McGeachy, A. M. & Weissman, J. S. The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nat. Protoc.* 7, 1534–1550 (2012).

35. Guo, H., Ingolia, N. T., Weissman, J. S. & Bartel, D. P. Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature* 466, 835–840 (2010).

36. Chung, B. Y. et al. The use of duplex-specific nuclease in ribosomal profiling and a user-friendly software package for Ribo-seq data analysis. *RNA* 21, 1731–1745 (2015).

37. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10, R25 (2009).

38. Butterworth, B. E. & Rueckert, R. R. Gene order of encephalomyocarditis virus as determined by studies with pactamycin. *J. Virol.* 9, 823–828 (1972).

39. Artieri, C. G. & Fraser, H. B. Accounting for biases in riboprofiling data indicates a major role for proline in stalling translation. *Genome Res.* 24, 453–463 (2014).

40. Naphine, S. et al. Expression of the VP2 protein of murine norovirus by a translation termination-reinitiation strategy. *PLoS ONE* 4, e8590 (2009).

41. Beach, D. L., Keane, J. D. in *Methods in Molecular Biology*. (ed. Wilusz, J.) Vol. 419 (Post-Transcriptional Gene Regulation, Humana Press, 2008).

42. Harper, S. & Speicher, D. W. Purification of proteins fused to glutathione S-transferase. *Methods Mol. Biol.* 681, 259–280 (2011).

43. Powell, M. L., Naphine, S., Jackson, R. J., Brierley, I. & Brown, T. D. Characterization of the termination-reinitiation strategy employed in the expression of influenza B virus BM2 protein. *RNA* 14, 2394–2406 (2008).