open
Biology

Ribosomal frameshifting used in influenza A virus expression occurs within the sequence UCC_UUU_CGU and is in the +1 direction

A. E. Firth, B. W. Jagger, H. M. Wise, C. C. Nelson, K. Parsawar, N. M. Wills, S. Napthine, J. K. Taubenberger, P. Digard and J. F. Atkins

References	**This article cites 58 articles, 30 of which can be accessed free** http://rsob.royalsocietypublishing.org/content/2/10/120109.full.html#ref-list-1
Subject collections	Articles on similar topics can be found in the following collections microbiology (11 articles) molecular biology (27 articles)
Email alerting service	Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click **here**

open Biology

Research

Author for correspondence:
J. F. Atkins
e-mail: j.atkins@ucc.ie

Ribosomal frameshifting used in influenza A virus expression occurs within the sequence UCC_UUU_CGU and is in the +1 direction

A. E. Firth[1,†], B. W. Jagger[1,2,†], H. M. Wise[1,3], C. C. Nelson[4], K. Parsawar[4], N. M. Wills[5], S. Napthine[1], J. K. Taubenberger[2], P. Digard[1,3] and J. F. Atkins[6]

[1]Division of Virology, Department of Pathology, University of Cambridge, Cambridge CB2 1QP, UK
[2]Viral Pathogenesis and Evolution Section, Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA
[3]The Roslin Institute, University of Edinburgh, Easter Bush, Midlothian EH25 9RG, UK
[4]Mass Spectrometry and Proteomics Core Facility, and [5]Department of Human Genetics, University of Utah, Salt Lake City, UT 84112, USA
[6]BioSciences Institute, University College Cork, Cork, Republic of Ireland

1. Summary

Programmed ribosomal frameshifting is used in the expression of many virus genes and some cellular genes. In eukaryotic systems, the most well-characterized mechanism involves −1 tandem tRNA slippage on an X_XXY_YYZ motif. By contrast, the mechanisms involved in programmed +1 (or −2) slippage are more varied and often poorly characterized. Recently, a novel gene, PA-X, was discovered in influenza A virus and found to be expressed via a shift to the +1 reading frame. Here, we identify, by mass spectrometric analysis, both the site (UCC_UUU_CGU) and direction (+1) of the frameshifting that is involved in PA-X expression. Related sites are identified in other virus genes that have previously been proposed to be expressed via +1 frameshifting. As these viruses infect insects (chronic bee paralysis virus), plants (fijiviruses and amalgamaviruses) and vertebrates (influenza A virus), such motifs may form a new class of +1 frameshift-inducing sequences that are active in diverse eukaryotes.

2. Introduction

During translation, shifts in reading register can occur to either alternative frame. The most widely known frameshifting mechanism involves shifting to the −1 frame. In part, this is because of the relatively well-defined nature of the most commonly used shift site motif that allows two adjacent tRNAs to re-pair to mRNA in the −1 frame, and in part due to the prominence of the viruses and other mobile elements that use this type of frameshift. The other reading frame can be accessed by either a −2 or a +1 frameshift event, with

[†]These authors contributed equally to this study.

the product of the former having an extra amino acid encoded by the shift site sequence relative to the latter.

In the majority of bacteria, frameshifting to the +1 frame is used as a sensor and effector of an autoregulatory circuit for the expression of release factor 2 [1,2]. In animals and fungi, such frameshifting is widely used to regulate expression of antizyme, the negative regulator of cellular polyamine levels [3,4]. In both cases, protein sequencing has shown that the shift is +1. Interestingly, however, although the mammalian antizyme 1 frameshifting signals exclusively drive +1 frameshifting in mammalian cells, they induce both +1 and −2 frameshifting when a cassette containing them is expressed in *Schizosaccharomyces pombe*, and −2 frameshifting when expressed in *Saccharomyces cerevisiae* [5]. In this system, the ratio of −2 to +1 is alterable depending on the distance of a 3′-adjacent stimulatory pseudoknot structure from the shift site [6]. Similarly, frameshifting on the HIV shift site U_UUU_UUA, which is normally −1, can be altered to −2 by varying the distance to the 3′ stimulatory element [7]. The only known natural case of programmed −2 frameshifting occurs during the expression of the gpGT tail assembly protein of phage Mu, where the efficiency of frameshifting is estimated to be about 2.2 per cent [8,9]. Protein sequencing has also been used to determine that +1 frameshifting is used in the expression of the *Tsh* gene of several *Listeria* phages and *Bacillus subtilis* SPP1 phage, besides *Escherichia coli yepP*, and the *pol* gene of the *S. cerevisiae* retrotransposons Ty1 and Ty3 [10–14]. Given similar sequences as in Ty1, the frameshifting used in decoding the mRNAs for actin filament binding protein ABP140 and telomere component EST3 is also expected to be +1 [15–18].

Frameshifting, probably in the +1 direction, has also been reported in mitochondria from several diverse species, although functionally different cases of frameshifting used in human mitochondria are −1 [19–22]. Peptide analysis has confirmed shifting to the +1 frame in one of the significant number of *Euplotes* genes that use such frameshifting, but the transframe-encoded peptide that would demonstrate the nature of the shift remains elusive [23,24]. Further work is also required on the early identified case involving the RNA phage MS2 coat lysis hybrid [25]. Low-efficiency cases are especially challenging—for instance, that of the clinically relevant shifting to the +1 frame that is seen in some cases of drug-resistant herpes simplex virus [26–28]. As a test case, even very low levels of the resulting frameshift product were shown to be able to function as an epitope for stimulation of CD8+ T cells [29].

Recently, Jagger *et al*. identified a novel coding ORF (X) in influenza A virus [30]. The X ORF is translated as a transframe fusion (PA-X) with the N-terminal domain of the PA protein (figure 1a). PA is a component of the viral polymerase, and the N-terminal domain carries an endonuclease activity that, as part of PA, cleaves capped RNA fragments from cellular pre-mRNAs to act as primers for viral transcription [31]. As PA-X, however, the N-terminal domain appears to play a role in host cell shut-off, presumably by cleaving host mRNAs. PA-X expression depends on ribosomal frameshifting into the +1 frame, and comparative sequence analysis suggests that the frameshifting occurs within a highly conserved UCC_UUU_CGU sequence at the 5′ end of the X ORF (underscores separate zero-frame, i.e. PA, codons) [30]. However, the exact site and direction of frameshifting was not determined. Here, we identify the nature of the shift

to the +1 frame in PA-X expression. The results highlight the coding versatility of the sequence UCC_UUU_CGU, with expression relevance for genomes (both viral and cellular) less well studied than influenza A virus.

3. Results and discussion

The efficiency of frameshifting at the PA-X shift site was previously estimated by translating reporter constructs in rabbit reticulocyte lysates and found to be around 1.3 per cent [30]. When the frameshift cassette was fused into a dual luciferase reporter construct and expressed in tissue culture cells (see §4), comparably low frameshifting efficiencies (namely $0.74 \pm 0.13\%$) were measured. Owing to the low levels involved and the lack of a suitably sensitive antibody to the common N-terminal domain of PA and PA-X, we have not been able to directly measure the frameshifting efficiency in the context of viral infection. Because PA-X is expressed at very low levels during virus infection, we were not able to isolate sufficient quantities from virus-infected cells for mass spectrometric analysis despite multiple attempts. Thus, in order to determine the precise site and direction of frameshifting, we used a construct in which an ORF-encoding green fluorescent protein (GFP) was fused in-frame to the 3′ end of the X ORF (figure 1b). Frameshift expression of the construct would result in the transframe fusion PA-X-GFP, which could be affinity-purified on GFP-TRAP beads, while non-frameshift expression would result in a product that does not contain GFP. The construct was expressed in 293T cells, and PA-X-GFP was affinity-purified from cell lysates and resolved by SDS-PAGE. An in-frame control, in which the predicted shift site UCC_UUU_CGU_C was mutated to UCC_UUU_GUC to force expression of PA-X-GFP, was also prepared to show the approximate size at which the frameshift protein should migrate in gels. The wild-type construct produced a specific band migrating at the expected size for PA-X-GFP. A gel slice containing this protein was excised, digested with trypsin, and the resulting peptides were analysed by nano-liquid chromatography tandem mass spectrometry (nano-LC/MS/MS).

Eight separate PA-X-GFP tryptic peptides were identified, including peptides encoded both upstream and downstream of the shift site (figure 1c; two of the peptides have overlapping sequence). Importantly, a peptide spanning the shift site itself was identified (figure 1d). This peptide, GLWDSFVSPR, defines the shift site (UCC_UUU_CGU) and direction (+1) of frameshifting (figure 1e). Molecular ions for GLWDSFVSPR were identified both with and without oxidation at the tryptophan, providing further support for the sequence assignment. No peptide compatible with −2 frameshifting was detected. Formally, the peptide GLWDSFVSPR is compatible with three different models for frameshifting: (i) +1 slippage with UUU in the P-site and an empty A-site; (ii) +1 slippage with UCC in the P-site and an empty A-site; and (iii) tandem +1 slippage with UCC in the P-site and UUU in the A-site. However, consideration of the potential for codon : anticodon re-pairings favours model (i). Both UUU and UUC are translated by a single tRNA isoacceptor whose anticodon, 3′-AAG-5′, has a higher affinity for UUC in the +1 frame than for the zero-frame UUU [32]. By contrast, UCC is expected to be generally decoded by the serine tRNA with anticodon 3′-AGI-5′ (I, inosine), but whether it is decoded
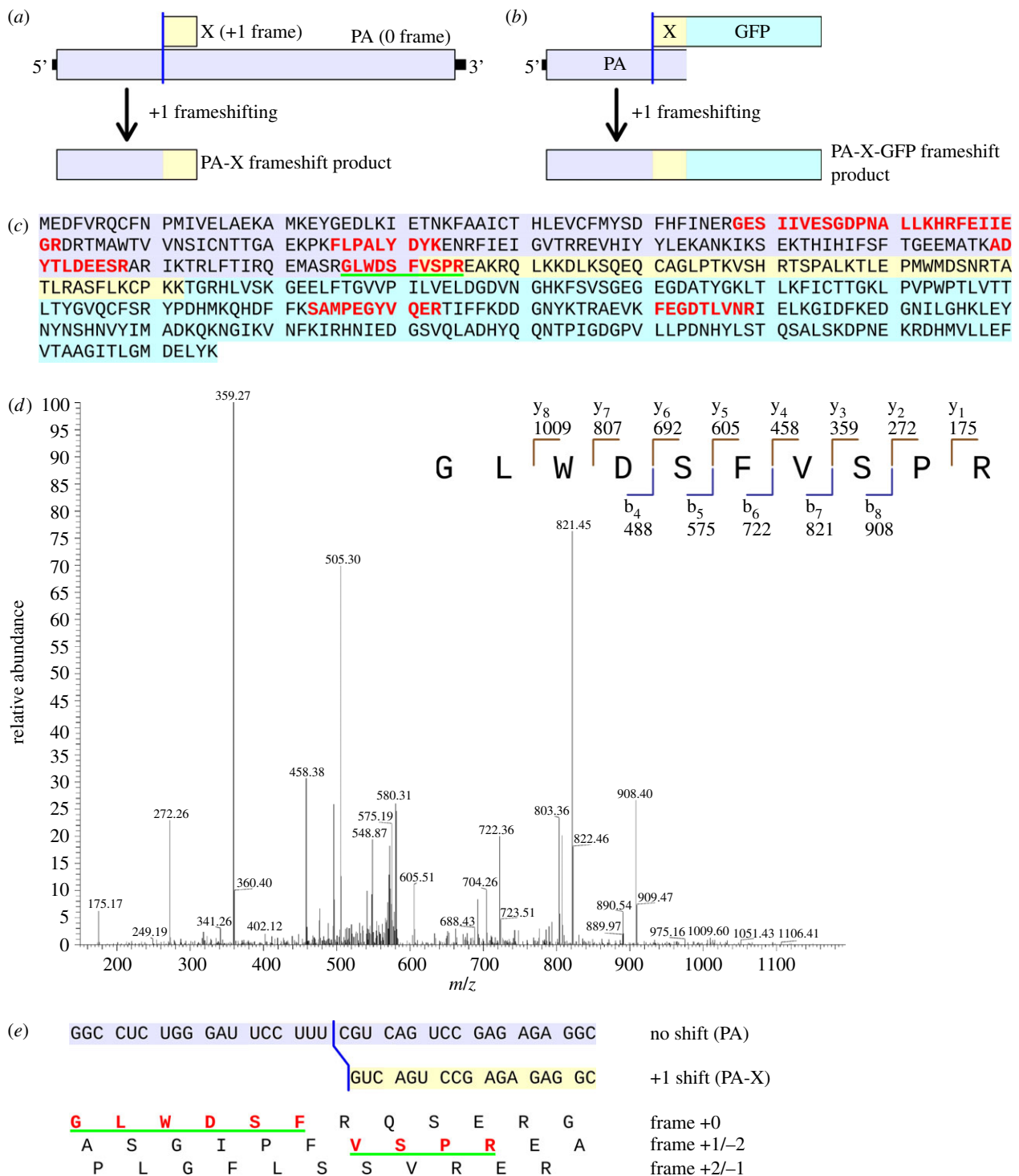
Figure 1. Mass spectrometric analysis of the PA-X-GFP frameshift fusion protein. (*a*) Translation map of influenza A virus segment 3 showing full-length PA and the transframe fusion PA-X that comprises the N-terminal domain of PA fused to a C-terminal tail encoded by the +1 reading frame. (*b*) Map of the construct used to purify the product of frameshifting on the PA-X frameshift cassette. (*c*) Complete amino acid sequence of PA-X-GFP. Amino acids encoded by the zero-frame are highlighted in mauve; amino acids encoded by the +1 frame are highlighted in pale yellow (X) or cyan (GFP). The eight peptides identified by mass spectrometry are indicated in red (note that the sequence GES…EGR corresponds to three detected peptides GES…LLK, HRF…EGR and FEI…EGR). The peptide spanning the frameshift site is underlined in green. (*d*) MS/MS fragmentation spectrum of the shift site peptide GLWDSFVSPR. The inset shows the peptide sequence with 'b-' and 'y'-type fragment ions that strongly support the shift site peptide identified in the nano-LC/MS/MS analysis. Several additional fragment ions, corresponding to H$_2$O losses from b and y series ions and doubly charged fragment ions, are also present in the spectrum to further support the sequence (assignments not labelled in the figure). (*e*) Nucleotide sequence in the vicinity of the frameshift site UCC_UUU_CGU, with conceptual amino acid translations in all three reading frames. The product of +1 frameshifting is indicated in red. The green-underlined peptide, which spans the shift site, is compatible with +1, but not −2, frameshifting.

by 3'-AGI-5' or a different serine tRNA when frameshifting occurs, re-pairing to CCU in the +1 frame would involve a mismatch at the first nucleotide position. Moreover, previous experiments showed that mutating UCC to AGC, GGG, CCC or AAA reduced but did not abolish frameshifting,

while mutating UUU_CGU to UUC_AGA (with an appropriately positioned 3' stop codon to prevent non-specific frameshifting elsewhere within the overlap region) knocked out frameshifting [30]. These results are consistent with P-site slippage on UUU_C but argue against P-site slippage

Table 1. Arginine codon usage frequencies (per 1000 codons) in selected organisms.

	Escherichia coli	human	chicken	bee	rice	*Arabidopsis*
AGA	2.9	12.2	12.2	22.0	10.5	19.0
AGG	1.8	12.0	11.7	9.1	16.0	11.0
CGU	20.2	4.5	5.4	10.5	7.2	9.0
CGC	20.8	10.4	10.4	5.1	16.1	3.8
CGA	3.8	6.2	5.3	11.4	6.4	6.3
CGG	6.2	11.4	9.7	4.1	13.4	4.9

on UCC_U, although a low level of slippage on UCC_U cannot be ruled out. Interestingly, a UCC_U tetranucleotide is the site of +1 frameshifting in antizyme expression, although here frameshifting is stimulated, in part, by the presence of a stop codon in the A-site (a role that is unlikely to be substituted by a UUU codon in the A-site) [3].

In other cases of +1 frameshifting, such as in bacteria and yeast, frameshifting is stimulated in part by a slowly decoded A-site codon such as a stop codon or codon whose cognate tRNA is limiting [1,33,34]. At the influenza PA-X shift site, P-site slippage on the UUU_C tetranucleotide may be stimulated by the rare CGU codon in the A-site (CGU is one of the most seldom-used codons in the genomes of mammals and birds—the host species of influenza A virus; table 1 [35]). In support of this, mutating the CGU to the more commonly used arginine codon, CGG, reduced frameshifting by 50 per cent [30]. However, CGU and the more abundantly used codon CGC are expected to be decoded by the same tRNA isoacceptor with anticodon 3′-GCI-5′, and this tRNA species is not obviously limiting in mammals and birds [36,37]. Thus, the role and mode of action of the A-site codon remains uncertain, and conservation of CGU may in part be driven by constraints on the encoded amino acid sequence in the overlapping +1 reading frame.

The role of UCC in the E-site also remains uncertain. In analyses of codon usage in PA, it was observed that the motif UCC_UUU_CGU is extremely highly conserved at the 5′ end of the influenza A virus X ORF, despite the fact that five other codons could potentially be used to encode the serine [30,38]. Moreover, mutating the UCC codon to AGC (serine) or to GGG, CCC or AAA resulted in a 40 to 70 per cent reduction in the frameshifting efficiency [30]. This suggests that UCC plays an important stimulatory role in the E-site. Earlier *in vivo* work on E-site influence (independent of amino acid identity) on stop codon readthrough implies that interactions at that site influence competition for A-site acceptance, but whether this influence acts via the P-site merits investigation [39,40]. Notwithstanding complications due to an interaction with rRNA during bacterial release factor 2 +1 frameshifting, there is evidence in that case for the identity of the E-site codon having an effect on +1 frameshifting. This has been proposed to relate to the speed at which the E-site tRNA is released, with weaker codon : anticodon duplexes being associated with higher levels of frameshifting [41–44]. In an *E. coli* cell-free system, even partially mismatched P-site codon : anticodon interactions, which can be augmented by E-site mismatches, trigger retrospective editing and so influence events in the A-site [45]. A counterpart post-peptide bond effect has not been detected in *S. cerevisiae*, but may exist and involve currently unidentified factors [46,47]. An E-site effect on +1 frameshifting could potentially be influenced

by the E-site tRNAs in a proportion of translating ribosomes being near-cognate rather than the standard cognate tRNA. The proposal of an allosteric relationship between release of deacylated tRNA from the E-site being coupled to aminoacyl-tRNA acceptance in the A-site [44] has drawn much criticism [48–51]. On its own, the observed E-site influence on +1 frameshifting could be interpreted as it acting via an effect on the length of the A-site pause that affects the probability of P-site realignment, but a direct effect on P-site codon : anticodon interaction, or rather on the translocating complex, seems more likely.

More generally, one might predict a class of +1 frameshift stimulators that comprise a UUU_C P-site slippery sequence and a restricted choice of A- and E-site codons. In eukaryote-infecting viruses, frameshifting by +1 nt has been predicted as the expression mechanism for non-5′-proximal ORFs in the closteroviruses (RdRp), leishmania RNA virus 1 (RdRp), chronic bee paralysis virus and the related Lake Sinai viruses 1 and 2 (RdRp), plant-infecting fijiviruses (Family Reoviridae; P5-2) and members of the proposed family *Amalgamaviridae* of plant viruses (RdRp) (reviewed in [52]). However, in most of these species, the site of frameshifting remained elusive. Characterization of the influenza virus frameshift site now suggests the site of +1 frameshifting in several of these viruses (figure 2). Several of these shift sites are also well supported by comparative genomic analysis [53]. Interestingly, these putative shift sites all seem to show a preference for A-site CGN codons, as opposed to other CNN codons. As in PA-X expression, it is likely that the efficiency of frameshifting at such sites is low. However, these levels may be completely compatible with the expression level requirements of some viruses (cf. −1 frameshifting for polymerase expression in *S. cerevisiae* totivirus L-A, where the ratio of Gag-Pol to Gag in the virion is of order 1–2% and, correspondingly, the frameshifting efficiency is around 1.8%) [54]. Whether similar motifs are functionally used for cellular gene expression remains to be seen.

4. Methods

4.1. Dual luciferase reporter constructs and assays

Sequences encompassing the frameshift site (97 nt 5′+ UCC_UUU_CGU+100 nt 3′) were generated using overlapping synthetic oligonucleotides and cloned into pDluc, a derivative of the dual luciferase reporter p2luc vector [55,56]. The 3′ firefly luciferase ORF is in the +1 frame relative to the 5′ renilla luciferase ORF, so that frameshifting within the inserted sequence results in a fusion of

```
(a) AY029521   2393-UCA UCA AAC CAG GCG UAU UCU UUU CGA AAC UAU AUC CGA UCC ACU AGA   FDV
    AY607587   2401-CUA CCU GGU AAA CAU AAA UCU UUU CGA AGC UUC CUA CGA UCU UCU CGA   MRCV
    AJ409147   2401-AGU UCC AGU CGA AGU AAG UCU UUU CGU UCC UAC UUA CGA UCU UCU CGA   RBSDV
    AY144569   2401-AGU UCC AGC CGA AGU AAG UCU UUU CGU UCC UAC UUA CGA UCU UCU CGA   RBSDV
    FN563987   2404-GGA ACC AGC AGA AAC AAG UCU UUU CGU UCC UAU UUA CGA UCU UCU CGA   SRBSDV
    FN563993   2404-GGA ACU AGC AAA AAC AAG UCU UUU CGU UCC UAU UUA CGA UCU UCU CGA   SRBSDV
    HM585275   2404-GGA ACU AGC AGA AAC AAG UCU UUU CGU UCC UAU UUA CGA UCU UCU CGA   SRBSDV
                  *             *       * *** *** **      *   *     *   *** **    ** **

(b) EU122229   1612---- --- --- --- --- AAG CCC UUU CGU GAC GGG CGC CCG AUA CUU GAU   CBPV
    EU122231   1603---- --- --- --- --- AAG CCC UUU CGU GAC GGG CGC CCG AUA CUU GAU   CBPV
    HQ871931   1758-ACC ACA UUC AAU CUG GAG CGA UUU CGU GCU GAG GCC UCG ACU GUU GGU   LSV1
    HQ888865   1698-ACA GUC UUC AAC CUU GAG CGA UUU CGU GCU GAG GCC ACG GUU GUU GGU   LSV2
                  **     *    *** *** *   * *     *   **      ** *  *

(c) HM029246    938-GCC AGG GAG AAG AUU CAG UCU UUU CGU GAC UCA GGA AAC CAG GGA GGA   BBLV

    HQ128706    971-GAA CCC AAG AAG GCA GGG ACU UUU CGC AGC CUC UUG GCA GUC CUC GGU   RhVA
    EU371896    977-GAG AAG AAG CAG GCG GGG ACU UUU CGU AAC CUC GUG GUG GCC GGA UGU   VCVM
                  **       *** ** ** *** *** *** ** * * *** ** * * *       **
```

Figure 2. Predicted sites of 'PAX-like' +1 frameshifting in (a) fijiviruses, (b) chronic bee paralysis and Lake Sinai viruses, and (c) amalgamaviruses. FDV, Fiji disease virus; MRCV, mal de Rio Cuarto virus; RBSDV, rice black-streaked dwarf virus; SRBSDV, southern rice black-streaked dwarf virus; CBPV, chronic bee paralysis virus; LSV, Lake Sinai virus; BBLV, blueberry latent virus; RhVA, rhododendron virus A; VCVM, *Vicia* cryptic virus M. In all cases, the predicted shift site occurs near the 5′ end of the overlap region between the zero-frame and +1 frame ORFs. Predicted shift sites are highlighted in blue. Dashes in CBPV indicate alignment gaps. Spaces separate zero-frame codons. Note that, downstream of the shift site, the sequences are predicted to be coding in both the zero and +1 frames, and this generally corresponds to enhanced conservation at the nucleotide level. The amalgamavirus sequences are highly divergent, and the precise alignment between BBLV and RhVA+VCVM is ambiguous in this region. GenBank accession numbers, and sequence coordinates of 5′ terminal nucleotides, are indicated at left.

both ORFs. An in-frame control, which was identical except that the UCC_UUU_CGU_C shift site was mutated to UCC_UUC_GUC, was also constructed. All constructs were verified by DNA sequencing. Frameshift assays were performed as described previously [55,57]. Frameshift efficiencies were calculated as (firefly activity/renilla activity) for the frameshift sequence normalized by (firefly activity/ renilla activity) for the in-frame control sequence. Means and standard errors were calculated based on four to six independent transfections. Owing to the low frameshifting efficiencies involved, a low level of background firefly activity (e.g. owing to cryptic splice sites, cryptic promoters, degraded transcripts or non-specific IRES activity) was a potential issue. To control for this, firefly and renilla activities were also measured for a corresponding shift-site mutant sequence (UUU_CGU mutated to UUC_AGA), and the ratio was subtracted from the ratio measured for the WT sequence. (It should be noted that independent initiation of the downstream reporter is not an issue for the previous frameshift efficiency measurements in rabbit reticulocyte lysates, where radiolabelled translation products could be visualized via SDS-PAGE.)

4.2. Protein purification

To create the PA-X-GFP expression construct, the nucleotide sequence corresponding to the coding region of PA-X, minus the X-ORF stop codon, was amplified from a A/ Brevig Mission/1/1918 (H1N1) segment 3 reverse genetics plasmid [58] and cloned into pEGFP-N1 using standard techniques (forward primer 5′-GCCACCGGTACCATGGAA GACTTTGTGCGACAATG-3′; reverse primer 5′-GCCACC ACCGGTCTTCTTTGGACATTTGAGAAAGC-3′). To avoid PA-X auto-repressing its own synthesis, the PA endonuclease active site was inactivated via the mutation D108A [30]. The GFP-initiating ATG was also mutagenized (ATG to TG) to bring the downstream GFP ORF in-frame with the +1-frameshifted X-ORF and to prevent downstream GFP initiation

(forward primer 5′-CCGGTCGCCACCTGGTGAGCAAGG-3′; reverse primer 5′-CCTTGCTCACCAGGTGGCGACCGG-3′). For the in-frame control construct, site-directed mutagenesis was used to delete the cytosine that is skipped during frameshifting, using standard techniques. Constructs were transfected into 293T cells using Lipofectamine 2000 (Invitrogen), according to the manufacturer's instructions. After incubation for 48 h, cells were lysed and GFP-TRAP-A purification (Chromotek) was performed, as previously described [59]. The GFP-TRAP bound fraction was resolved by SDS-PAGE, and polypeptides were visualized by silver staining.

4.3. Mass spectrometric analysis

Gel slices containing proteins of interest were excised, digested with trypsin, and analysed by nano-LC/MS/MS. All mass spectra were acquired with an LTQ-FT instrument (ThermoElectron). nano-LC with nano-electrospray was used with a 75 µm ID column (C18) and an acetonitrile gradient (0.1% formic acid). Primary mass spectra of peptide molecular ions, primarily observed at +2 charge states, were obtained in the FT-ICR (Fourier transform ion cyclotron resonance) part of the instrument. All peptide masses assigned were better than 2 ppm mass error compared with theoretical values. Both oxidized (i.e. addition of O, occurring at methionine, tryptophan or histidine) and non-oxidized forms were identified for many peptides. Oxidation of peptides is a common occurrence observed during ionization with electrospray, but oxidation can also be present as a post-translational event. Peptide sequence information was acquired using MS/MS with the ion-trap part of the LTQ-FT instrument using collision-induced dissociation fragmentation of selected peptide masses. Peptides were assigned based on combined evidence of the molecular ions and MS/MS sequence. Searches of custom sequence databases were performed with MASCOT [60], using strict parameters to generate high-confidence assignments, and, in addition, all primary and MS/MS data were reviewed manually for accuracy.

5. Acknowledgements

References

1. Craigen WJ, Caskey CT. 1986 Expression of peptide chain release factor 2 requires high efficiency frameshift. *Nature* **322**, 273–275. (doi:10.1038/322273a0)

2. Bekaert M, Atkins JF, Baranov PV. 2006 ARFA: a program for annotating bacterial release factor genes, including prediction of programmed ribosomal frameshifting. *Bioinformatics* **22**, 2463–2465. (doi:10.1093/bioinformatics/btl430)

3. Ivanov IP, Atkins JF. 2007 Ribosomal frameshifting in decoding antizyme mRNAs from yeast and protists to humans: close to 300 cases reveal remarkable diversity despite underlying conservation. *Nucleic Acids Res.* **35**, 1842–1858. (doi:10.1093/nar/gkm035)

4. Kurian L, Palanimurugan R, Gödderz D, Dohmen RJ. 2011 Polyamine sensing by nascent ornithine decarboxylase antizyme stimulates decoding of its mRNA. *Nature* **477**, 490–494. (doi:10.1038/nature10393)

5. Ivanov IP, Gesteland RF, Matsufuji S, Atkins JF. 1998 Programmed frameshifting in the synthesis of mammalian antizyme is +1 in mammals, predominantly +1 in fission yeast, but −2 in budding yeast. *RNA* **4**, 1230–1238. (doi:10.1017/S1355838298980864)

6. Matsufuji S, Matsufuji T, Wills NM, Gesteland RF, Atkins JF. 1996 Reading two bases twice: mammalian antizyme frameshifting in yeast. *EMBO J.* **15**, 1360–1370. (doi:10.1006/geno.1996.0601)

7. Lin Z, Gilbert RJ, Brierley I. 2012 Spacer-length dependence of programmed −1 or −2 ribosomal frameshifting on a U6A heptamer supports a role for messenger RNA (mRNA) tension in frameshifting. *Nucleic Acids Res.* **40**, 8674–8689. (doi:10.1093/nar/gks629)

8. Xu J, Hendrix RW, Duda RL. 2004 Conserved translational frameshift in dsDNA bacteriophage tail assembly genes. *Mol. Cell* **16**, 11–21. (doi:10.1016/j.molcel.2004.09.006)

9. Baranov PV, Fayet O, Hendrix RW, Atkins JF. 2006 Recoding in bacteriophages and bacterial IS elements. *Trends Genet.* **22**, 174–181. (doi:10.1016/j.tig.2006.01.005)

10. Dorscht J, Klumpp J, Bielmann R, Schmelcher M, Born Y, Zimmer M, Calendar R, Loessner MJ. 2009 Comparative genome analysis of *Listeria* bacteriophages reveals extensive mosaicism, programmed translational frameshifting, and a novel prophage insertion site. *J. Bacteriol.* **191**, 7206–7215. (doi:10.1128/JB.01041-09)

11. Auzat I, Dröge A, Weise F, Lurz R, Tavares P. 2008 Origin and function of the two major tail proteins of bacteriophage SPP1. *Mol. Microbiol.* **70**, 557–569. (doi:10.1111/j.1365-2958.2008.06435.x)

12. Liao PY, Choi YS, Lee KH. 2009 FSscan: a mechanism-based program to identify +1 ribosomal frameshift hotspots. *Nucleic Acids Res.* **37**, 7302–7311. (doi:10.1093/nar/gkp796)

13. Belcourt MF, Farabaugh PJ. 1990 Ribosomal frameshifting in the yeast retrotransposon Ty: tRNAs induce slippage on a 7 nucleotide minimal site. *Cell* **62**, 339–352. (doi:10.1016/0092-8674(90)90371-K)

14. Farabaugh PJ, Zhao H, Vimaladithan A. 1993 A novel programed frameshift expresses the POL3 gene of retrotransposon Ty3 of yeast: frameshifting without tRNA slippage. *Cell* **74**, 93–103. (doi:10.1016/0092-8674(93)90297-4)

15. Asakura T *et al.* 1998 Isolation and characterization of a novel actin filament-binding protein from *Saccharomyces cerevisiae*. *Oncogene* **16**, 121–130. (doi:10.1038/sj.onc.1201487)

16. Farabaugh PJ, Kramer E, Vallabhaneni H, Raman A. 2006 Evolution of +1 programmed frameshifting signals and frameshift-regulating tRNAs in the order *Saccharomycetales*. *J. Mol. Evol.* **63**, 545–561. (doi:10.1007/s00239-005-0311-0)

17. Morris DK, Lundblad V. 1997 Programmed translational frameshifting in a gene required for yeast telomere replication. *Curr. Biol.* **7**, 969–976. (doi:10.1016/S0960-9822(06)00416-7)

18. Taliaferro D, Farabaugh PJ. 2007 An mRNA sequence derived from the yeast EST3 gene stimulates programmed +1 translational frameshifting. *RNA* **13**, 606–613. (doi:10.1261/rna.412707)

19. Russell RD, Beckenbach AT. 2008 Recoding of translation in turtle mitochondrial genomes: programmed frameshift mutations and evidence of a modified genetic code. *J. Mol. Evol.* **67**, 682–695. (doi:10.1007/s00239-008-9179-0)

20. Masuda I, Matsuzaki M, Kita K. 2010 Extensive frameshift at all AGG and CCC codons in the mitochondrial cytochrome c oxidase subunit 1 gene of *Perkinsus marinus* (Alveolata; Dinoflagellata). *Nucleic Acids Res.* **38**, 6186–6194. (doi:10.1093/nar/gkq449)

21. Oudot-Le Secq MP, Green BR. 2011 Complex repeat structures and novel features in the mitochondrial genomes of the diatoms *Phaeodactylum tricornutum* and *Thalassiosira pseudonana*. *Gene* **476**, 20–26. (doi:10.1016/j.gene.2011.02.001)

22. Temperley R, Richter R, Dennerlein S, Lightowlers RN, Chrzanowska-Lightowlers ZM. 2010 Hungry codons promote frameshifting in human mitochondrial ribosomes. *Science* **327**, 301. (doi:10.1126/science.1180674)

23. Aigner S, Lingner J, Goodrich KJ, Grosshans CA, Shevchenko A, Mann M, Cech TR. 2000 Euplotes telomerase contains an La motif protein produced by apparent translational frameshifting. *EMBO J.* **19**, 6230–6239. (doi:10.1093/emboj/19.22.6230)

24. Vallabhaneni H, Fan-Minogue H, Bedwell DM, Farabaugh PJ. 2009 Connection between stop codon reassignment and frequent use of shifty stop frameshifting. *RNA* **15**, 889–897. (doi:10.1261/rna.1508109)

25. Atkins JF, Gesteland RF, Reid BR, Anderson CW. 1979 Normal tRNAs promote frameshifting. *Cell* **18**, 1119–1131. (doi:10.1016/0092-8674(79)90225-3)

26. Hwang CB, Horsburgh B, Pelosi E, Roberts S, Digard P, Coen DM. 1994 A net +1 frameshift permits synthesis of thymidine kinase from a drug-resistant herpes simplex virus mutant. *Proc. Natl Acad. Sci. USA* **91**, 5461–5465. (doi:10.1073/pnas.91.12.5461)

27. Griffiths A. 2011 Slipping and sliding: frameshift mutations in herpes simplex virus thymidine kinase and drug-resistance. *Drug Resist. Updat.* **14**, 251–259. (doi:10.1016/j.drup.2011.08.003)

28. Pan D, Coen DM. 2012 Quantification and analysis of thymidine kinase expression from acyclovir-resistant G-string insertion and deletion mutants in herpes simplex virus-infected cells. *J. Virol.* **86**, 4518–4526. (doi:10.1128/JVI.06995-11)

29. Zook MB, Howard MT, Sinnathamby G, Atkins JF, Eisenlohr LC. 2006 Epitopes derived by incidental translational frameshifting give rise to a protective CTL response. *J. Immunol.* **176**, 6928–6934.

30. Jagger BW *et al.* 2012 An overlapping protein-coding region in influenza A virus segment 3 modulates the host response. *Science* **337**, 199–204. (doi:10.1126/science.1222213)

31. Boivin S, Cusack S, Ruigrok RW, Hart DJ. 2010 Influenza A virus polymerase: structural insights into replication and host adaptation mechanisms. *J. Biol. Chem.* **285**, 28 411–28 417. (doi:10.1074/jbc.R110.117531)

32. Eisinger J, Feuer B, Yamane T. 1971 Codon-anticodon binding in tRNA^{phe}. *Nat. New Biol.* **231**, 126–128. (doi:10.1038/231126a0)

33. Weiss RB, Dunn DM, Atkins JF, Gesteland RF. 1987 Slippery runs, shifty stops, backward steps, and forward hops: −2, −1, +1, +2, +5, and +6 ribosomal frameshifting. *Cold Spring Harb. Symp. Quant. Biol.* **52**, 687–693. (doi:10.1101/SQB.1987.052.01.078)

34. Farabaugh PJ. 2010 Programmed frameshifting in budding yeast. In *Recoding: expansion of decoding rules enriches gene expression* (eds JF Atkins, RF Gesteland), pp. 221–248. Heidelberg, Germany: Springer.

35. Nakamura Y, Gojobori T, Ikemura T. 2000 Codon usage tabulated from the international DNA sequence databases: status for the year 2000. *Nucleic Acids Res.* **28**, 292. (doi:10.1093/nar/28.1.292)

36. Grosjean H, de Crécy-Lagard V, Marck C. 2010 Deciphering synonymous codons in the three domains of life: co-evolution with specific tRNA modification enzymes. *FEBS Lett.* **584**, 252–264. (doi:10.1016/j.febslet.2009.11.052)

37. Chan PP, Lowe TM. 2009 GtRNAdb: a database of transfer RNA genes detected in genomic sequence. *Nucleic Acids Res.* **37**, D93–D97. (doi:10.1093/nar/gkn787)

38. Gog JR *et al*. 2007 Codon conservation in the influenza A virus genome defines RNA packaging signals. *Nucleic Acids Res.* **35**, 1897–1907. (doi:10.1093/nar/gkm087)

39. O'Connor M, Wills NM, Bossi L, Gesteland RF, Atkins JF. 1993 Functional tRNAs with altered 3′ ends. *EMBO J.* **12**, 2559–2566.

40. Sergiev PV, Lesnyak DV, Kiparisov SV, Burakovsky DE, Leonov AA, Bogdanov AA, Brimacombe R, Dontsova OA. 2005 Function of the ribosomal E-site: a mutagenesis study. *Nucleic Acids Res.* **33**, 6048–6056. (doi:10.1093/nar/gki910)

41. Baranov PV, Gesteland RF, Atkins JF. 2002 Recoding: translational bifurcations in gene expression. *Gene* **286**, 187–201. (doi:10.1016/S0378-1119(02)00423-7)

42. Sanders CL, Curran JF. 2007 Genetic analysis of the E-site during RF2 programmed frameshifting. *RNA* **13**, 1483–1491. (doi:10.1261/rna.638707)

43. Liao PY, Gupta P, Petrov AN, Dinman JD, Lee KH. 2008 A new kinetic model reveals the synergistic effect of E-, P- and A-sites on +1 ribosomal frameshifting. *Nucleic Acids Res.* **36**, 2619–2629. (doi:10.1093/nar/gkn100)

44. Pech M, Vesper O, Yamamoto H, Wilson DN, Nierhaus KH. 2010 The E site and its importance for improving accuracy and preventing frameshifts. In *Recoding: expansion of decoding rules enriches gene expression* (eds JF Atkins, RF Gesteland), pp. 345–364. Heidelberg, Germany: Springer.

45. Zaher HS, Green R. 2009 Quality control by the ribosome following peptide bond formation. *Nature* **457**, 161–166. (doi:10.1038/nature07582)

46. Eyler DE, Green R. 2011 Distinct response of yeast ribosomes to a miscoding event during translation. *RNA* **17**, 925–932. (doi:10.1261/rna.2623711)

47. Zaher HS, Green R. 2011 A primary role for release factor 3 in quality control during translation elongation in *Escherichia coli*. *Cell* **147**, 396–408. (doi:10.1016/j.cell.2011.08.045)

48. Semenkov YP, Rodnina MV, Wintermeyer W. 1996 The 'allosteric three-site model' of elongation cannot be confirmed in a well-defined ribosome system from *Escherichia coli*. *Proc. Natl Acad. Sci. USA*. **93**, 12 183–12 188. (doi:10.1073/pnas.93.22.12183)

49. Uemura S, Aitken CE, Korlach J, Flusberg BA, Turner SW, Puglisi JD. 2010 Real-time tRNA transit on single translating ribosomes at codon resolution. *Nature* **464**, 1012–1017. (doi:10.1038/nature08925)

50. Chen C, Stevens B, Kaur J, Smilansky Z, Cooperman BS, Goldman YE. 2011 Allosteric vs. spontaneous exit-site (E-site) tRNA dissociation early in protein synthesis. *Proc. Natl Acad. Sci. USA* **108**, 16 980–16 985. (doi:10.1073/pnas.1106999108)

51. Petropoulos AD, Green R. 2012 Further *in vitro* exploration fails to support the allosteric three-site model. *J. Biol. Chem.* **287**, 11 642–11 648. (doi:10.1074/jbc.C111.330068)

52. Firth AE, Brierley I. 2012 Non-canonical translation in RNA viruses. *J. Gen. Virol.* **93**, 1385–1409. (doi:10.1099/vir.0.042499-0)

53. Firth AE, Atkins JF. 2009 Analysis of the coding potential of the partially overlapping 3′ ORF in segment 5 of the plant fijiviruses. *Virol. J.* **17**, 32. (doi:10.1186/1743-422X-6-32)

54. Dinman JD, Icho T, Wickner RB. 1991 A −1 ribosomal frameshift in a double-stranded RNA virus of yeast forms a gag-pol fusion protein. *Proc. Natl Acad. Sci. USA* **88**, 174–178. (doi:10.1073/pnas.88.1.174)

55. Grentzmann G, Ingram JA, Kelly PJ, Gesteland RF, Atkins JF. 1998 A dual-luciferase reporter system for studying recoding signals. *RNA* **4**, 479–486. (doi:10.1017/S1355838298971576)

56. Fixsen SM, Howard MT. 2010 Processive selenocysteine incorporation during synthesis of eukaryotic selenoproteins. *J. Mol. Biol.* **399**, 385–396. (doi:10.1016/j.jmb.2010.04.033)

57. Baranov PV, Henderson CM, Anderson CB, Gesteland RF, Atkins JF, Howard MT. 2005 Programmed ribosomal frameshifting in decoding the SARS-CoV genome. *Virology* **332**, 498–510. (doi:10.1016/j.virol.2004.11.038)

58. Tumpey TM *et al*. 2005 Characterization of the reconstructed 1918 Spanish influenza pandemic virus. *Science* **310**, 77–80. (doi:10.1126/science.1119392)

59. Amorim MJ, Bruce EA, Read EK, Foeglein A, Mahen R, Stuart AD, Digard P. 2011 A Rab11- and microtubule-dependent mechanism for cytoplasmic transport of influenza A virus viral RNA. *J. Virol.* **85**, 4143–4156. (doi:10.1128/JVI.02606-10)

60. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. 1999 Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**, 3551–3567. (doi:10.1002/(SICI)1522-2683(19991201)20:18<3551::AID-ELPS3551>3.0.CO;2-2)