



Deposited via The University of York.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/244497/>

Version: Accepted Version

Article:

Betts, Jemma K, Jeffries, Cy M, Passchier, Tim C et al. (Accepted: 2026) A protein-dependent riboswitch activates ribosomal frameshifting in cardioviruses. *Molecular Cell*. ISSN: 1097-2765 (In Press)

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

A protein-dependent riboswitch activates ribosomal frameshifting in cardioviruses

Authors

Jemma K. Betts^{1,2,3}, Cy M. Jeffries^{4,5}, Tim C. Passchier^{1,2,3}, Hermia C. Y. Kung⁶, Sarah P. Graham^{7,8}, Mahmoud A.S. Abdelhamid^{9,10,11}, Jamieson A. L. Howard⁷, Timothy D. Craggs^{9,10,11}, Stephen C. Graham¹², Ian Brierley¹², Mark C. Leake^{2,3,7}, Steven D. Quinn^{2,7}, Chris H. Hill^{1,2,3}✉

Affiliations

1. York Structural Biology Laboratory, University of York, York, YO10 5DD, UK
2. York Biomedical Research Institute, University of York, York, YO10 5DD, UK
3. Department of Biology, University of York, York, YO10 5DD, UK
4. European Molecular Biology Laboratory (EMBL), Hamburg Unit, c/o Deutsches Elektronen-Synchrotron (DESY), Notkestrasse 85, 22607 Hamburg, Germany
5. Ilse Katz Institute for Nanoscale Science and Technology, Ben-Gurion University of the Negev, Be'er Sheva, 8410501, Israel
6. School of Natural Sciences, University of York, York, YO10 5DD, UK
7. School of Physics, Engineering & Technology, University of York, York, YO10 5DD, UK
8. Present address: Department of Biosciences, Durham University, Durham, DH1 3LE, UK
9. School of Mathematical and Physical Sciences, University of Sheffield, Sheffield, S3 7RH, UK
10. Centre for Single-molecule Biology, University of Sheffield, Sheffield, S10 2TN, UK
11. Exciting Instruments Ltd, Block 5, Level 9, Pennine Five Campus, Sheffield, S1 4WP, UK
12. Division of Virology, Department of Pathology, University of Cambridge, Cambridge, CB2 1QP, UK

✉ Lead author: chris.hill@york.ac.uk

Summary

Programmed -1 ribosomal frameshifting (PRF) is a translational control mechanism used by RNA viruses to regulate the relative abundance of proteins encoded in different reading frames. Cardioviruses exhibit the highest known PRF efficiency, with ~85% of ribosomes shifting into the -1 frame. This unusual event requires an interaction between the viral 2A protein and a stimulatory element in the RNA genome, but the basis for protein-dependence is unclear. To address this, here we investigate structure and dynamics of the PRF signal in Theiler's murine encephalitis virus (TMEV). By combining X-ray crystallography, small angle X-ray scattering (SAXS) and single-molecule fluorescence resonance energy transfer (smFRET), we show that 2A binding switches the RNA from a stem-loop conformation into a pseudoknot, and we demonstrate that pseudoknot formation is essential for efficient PRF *in vitro* and in cells. Together, these findings illustrate how the cardiovirus PRF element behaves as a protein-dependent riboswitch, defining the molecular mechanism by which frameshifting is conditionally activated.

Introduction

Programmed -1 ribosomal frameshifting (PRF) is a translational control mechanism that underpins the replication cycle of hundreds of RNA viruses (reviewed in¹⁻³), the collective impact of which are enormous. Examples include HIV-1 (*Lentivirus humimdef1*, ~1.3 million infections per year⁴), SARS-CoV-2 (*Betacoronavirus pandemicum*, ~760 million cases and ~6.9 million deaths⁵), Japanese encephalitis virus (JEV; *Orthoflavivirus japonicum*, ~100,000 infections and ~25,000 deaths per year⁶) and porcine reproductive and respiratory syndrome virus (PRRSV; *Betaarterivirus europensis*, ~\$664 million annual loss in the USA⁷).

During a -1 frameshift event, ribosomes are diverted one nucleotide backwards into an alternative reading frame at a specific position on the viral RNA, leading to the translational fusion of polypeptides encoded by the upstream '0 frame' and downstream '-1 frame' sequences^{8,9}. The frameshift efficiency is usually fixed, producing a ratio of 0 frame to -1 frame proteins that is optimal for efficient virus replication. Frameshifting is therefore crucial for viral fitness, and sequences directing PRF are highly conserved. Moreover, PRF is not widely used as a gene expression strategy by the host, making it an attractive target for antiviral therapeutics¹⁰⁻¹².

Frameshifting typically requires two elements. First, a heptameric 'slippery sequence' of the form X_XXY_YYZ (where X is any nucleotide, Y is A or U, and Z is any nucleotide except G) permits codon-anticodon base pairing in both 0 (X_XXY_YYZ) and -1 (XXX_YYY) frames. Secondly, a structured RNA 'stimulatory element' (usually a pseudoknot or stem-loop) located 5–9 nucleotides downstream acts as a mechanical barrier, resisting unwinding by ribosomal helicase proteins that encircle the mRNA entry channel^{13,14}. Ribosomes slow down or pause when they encounter the stimulatory element¹⁵⁻¹⁸, positioning the slippery sequence in the P and A sites, and causing tension to develop in the RNA¹⁹. Tandem slippage of tRNAs likely occurs during subsequent translocation attempts (reviewed in²⁰), characterised by prolonged rotated hybrid and chimeric states, impaired small subunit head swivel, abnormal elongation factor G (EF-G) function and slow release of E site tRNA^{21,22}. During this delay, tRNAs may spontaneously fluctuate between reading frames, reaching an equilibrium based on the relative free energies of pairing in the 0 and -1 frames^{22,23}.

RNA stimulatory elements present an intriguing paradox, with many — and often conflicting — features linked to their ability to promote frameshifting. Despite being protein-coding sequences, they must be stable enough to impede the ribosome, whilst also unwinding at physiologically achievable forces during translocation (~13 pN)²⁴. Early work suggested that thermodynamic stability of the stimulatory element would correlate with PRF efficiency²⁵, but this is an oversimplification. Subsequent structural work on diverse viral PRF pseudoknots revealed complex features hypothesised to be important, including minor groove triplex interactions between stem 1 (S1) and loop 3 (L3)²⁶⁻³⁰, unusual geometry such as base triples, quadruples or bending at the S1/S2 junction³¹⁻³³ and 'threading' of the 5' end in the SARS-CoV-2 pseudoknot³⁴⁻³⁶. Recently, a cryo-EM study revealed specific interactions between the SARS-CoV-2 pseudoknot and components of the 80S ribosomal helicase³⁷, providing evidence for the importance of specific RNA structures. Alternatively, it has been proposed that frameshifting is driven by the existence of multiple interconverting conformations that are sampled by a stimulatory element³⁸⁻⁴¹. These ideas can be difficult to reconcile with high-resolution structures of stimulatory elements, which are necessarily static snapshots of a single RNA conformer. This is further confounded by the inaccuracy of deep-learning-based RNA structure prediction tools⁴².

Another layer of complexity is presented by cellular and viral proteins, which can either inhibit⁴³⁻⁴⁵ or stimulate⁴⁶⁻⁴⁹ PRF. A striking case of the latter occurs in the cardioviruses (family

Picornaviridae), with 85% of ribosomes shifting into the -1 frame at a conserved G_GUU_UUU slippery sequence in Theiler's murine encephalitis virus (TMEV; *Cardiovirus theileri*)^{50–52} (**Figure 1A–C**). This is the highest known efficiency viral PRF event in nature, but is unusual for several reasons. First, it is entirely conditional on the presence of the viral 2A protein, which binds to an RNA stimulatory element located 13–14 nt downstream^{50,51}. Secondly, although this element is predicted to form a seven base pair stem-loop, none of the essential nucleotides reside within the stem. Instead, a pair of guanines in the flanking 5' region and a cytosine triplet in the loop are required for 2A binding and frameshifting⁵². Finally, the positioning of this element deviates significantly from the canonical 5–9 nt spacer — seemingly too far away to pause the ribosome over the slippery sequence. We previously hypothesised that the cardiovirus stimulatory element may instead form an RNA pseudoknot, which 2A selectively recognises^{52,53}. However, this does not adequately explain the protein-dependency of PRF, and structures of 2A proteins from TMEV and encephalomyocarditis virus (EMCV; *Cardiovirus rueckerti*) reveal no homology to known RNA-binding folds^{52,53}. Thus, the molecular mechanisms underpinning cardiovirus PRF remain elusive.

Here, we combine X-ray crystallography, SAXS and single-molecule fluorescence resonance energy transfer (smFRET) to examine the structure and dynamics of the TMEV stimulatory element in the presence and absence of 2A protein. We first determine a 1.9 Å structure of the 2A-RNA complex, revealing a compact RNA pseudoknot with a non-canonical two-stem architecture. We next demonstrate that the RNA stimulatory element behaves as a protein-dependent riboswitch, undergoing a dramatic conformational rearrangement from a seven base pair stem-loop into a pseudoknot upon 2A binding. Our structure explains the protein dependence of pseudoknot formation, and we dissect the sequence specificity of this unique molecular switch, *in vitro* and in cells. Together, this work defines the conformational repertoire of a protein-dependent riboswitch and demonstrates that — even though cardiovirus PRF elements are indeed conformationally plastic — one specific structure is likely the major determinant of PRF. We suggest that frameshifting signals in other viruses may be similarly regulated by as-yet-unidentified proteins or ligands, explaining some of the reported conformational heterogeneity.

Results

The PRF stimulatory element from TMEV is a protein-pseudoknot complex

To understand the molecular basis of PRF in the cardioviruses, we began with structural studies of the stimulatory element from TMEV. We determined the X-ray crystal structure of a complex between TMEV 2A protein (amino acids 1–133, corresponding to 923–1055 of the TMEV polyprotein, UniProt: [P08545](#)) and a minimal RNA stimulatory element (nucleotides 14–47 in **Figure 1B**, corresponding to 4257–4290 of the TMEV genome, GenBank: [NC_001366.1](#)). The structure was refined at 1.9 Å resolution (PDB: [9RVP](#), **Table 1**) revealing that the RNA adopts a pseudoknot conformation when bound to 2A protein (**Figure 1D** and **movie S1**). 2A adopts the same globular 'beta-shell' architecture as the *apo* structure⁵³ (**Figure S1A**, C_α backbone RMSD of 0.39 Å over 125 residues).

The RNA pseudoknot comprises two non-coaxially stacked stems in a side-by-side arrangement, with a tilt of ~55° between their helical axes. (**Figure 2** and **Figure S1B**). Stem 1 (S1) consists of three Watson-Crick G-C base pairs between GGxG and CCC triplet motifs, providing a structural basis to the previous biochemical characterisation^{49–53}. Interestingly, the three guanosine nucleotides (G17, G18 and G20) comprising S1 are not consecutive, and U19 is flipped out by S1 formation. This results in an unusual helical geometry: the first two G-

C base pairs (G17-C38 and G18-C37) exhibit a typical helical twist of $\sim 36.2^\circ$, but the final base pair (G20-C36) has a twist of $\sim 50.6^\circ$ relative to G18-C37 (**Figure S1C**), causing compression of the major groove.

Such geometry is unfavourable, but S1 is stabilised by the insertion of the 2A 'arginine loop' into the major groove (**Figure 2** and [movie S1](#)). The guanidinium groups of R87 and R85 are approximately coplanar with the second (G18-C37) and third (G20-C36) base pairs of S1 respectively, forming two 'arginine forks' — recurrent structural motifs in protein-RNA interactions^{54–56}. R87 stabilises the second base pair of S1 (G18-C37) via interactions between N η and the Hoogsteen edge of G18, and a hydrogen bond between N ϵ and a ribonucleotide phosphate oxygen of G17 (H η η P fork). R85 stabilises the third base pair of S1 (G20-C36) via interactions between N η 2 and N ϵ atoms and the Hoogsteen edge of G20, and between N η 2 and the ribose 2' hydroxyl of G18 (HS fork). On the other side of G20, the carbonyl oxygen of glycine 27 forms a hydrogen bond with N2, such that this guanine base is "clamped" between R85 and glycine 27. Comparison to the *apo* structure of TMEV 2A illustrates how glycine 27 moves 1.6 Å towards G20 while the guanidinium group of R85 rotates: we therefore refer to this as the "G-clamp" interaction (**Figure S1A**). This third base pair (G20-C36) additionally packs against the 2A beta-sheet, comprising R7, D9, F11 from β 1, H67 from β 4 and K24 from β 2, the latter of which forms a hydrogen bond with the ribose 2' hydroxyl of G20. Immediately prior to S1, the 5' end of the stimulatory RNA also interacts with the surface of 2A, with A15 stacking against W89 (**Figure 2**). Together, these interactions bury a surface of 634 Å². AlphaFold⁵⁷ was unable to predict this interface, instead suggesting interactions between 2A protein and S2 (**Figure S1D**).

Downstream of S1, loop 1 (L1) is a single nucleotide (C21) leading into stem 2 (S2), a 7 bp A-form helix (C22–A28; U39–G45) (**Figure 2**), consistent with previous bioinformatic predictions and structure probing experiments^{50,51}. S2 is connected to the 3' strand of S1 by loop 2 (L2, A29–U35, **Figure 2**). However, the electron density for L2 is poor (**Figure 2**) demonstrating that U31, A32 and A33 are flexible or disordered. Nevertheless, A29 and C30 stack against A28 from S2, and weak density suggests that A34 may stack against C38 from S1. Loop 3, a feature of some H-type pseudoknots⁵⁸, is entirely absent.

The RNA stimulatory element switches conformation upon 2A binding

We next examined to what extent our crystal structure reflects the conformation of 2A protein and RNA in solution. To do this, we analysed 2A protein, RNA, and the 2A-RNA complex separately by size exclusion chromatography coupled to small angle X-ray scattering (SEC-SAXS, **Figure S2A-F**, **Figure 3A-B** and **Table 2**). All samples were monodisperse, allowing determination of the radius of gyration (R_g), maximum particle dimension (D_{max}), and the calculation of multiple molecular envelopes and atomistic models describing the data (**Figure S2** and **Table 2**). The 2A protein sample (**Figure S2A**, $R_g = 1.89 \pm 0.05$ nm) matched the protein component of the crystal structure, with an ensemble of five models (**Figure S2B**, $\chi^2 = 0.92$ – 0.97) illustrating the conformational flexibility of disordered N- and C-termini. Similarly, the 2A-RNA complex sample (**Figure 3A** and **Figure S2E**, $R_g = 2.23 \pm 0.02$ nm) was in good agreement with the crystal structure (**Figure S2F**, five models, $\chi^2 = 1.11$ – 1.35), confirming that the pseudoknot-protein complex likely forms in solution. However, the scattering data from the RNA-only sample (**Figure S2C**, $R_g = 1.91 \pm 0.02$ nm) could not be explained by the pseudoknot component of the crystal structure (**Figure S2G**, $\chi^2 = 13.9$) and attempts to further refine the RNA pseudoknot to match the SAXS data were unsuccessful. Instead, these data are explained by a 7 bp stem-loop with a flexible unpaired 5' tail (**Figure 3B** and **Figure S2D**, nine

models, $\chi^2 = 1.11\text{--}1.20$), suggesting that the RNA may undergo a large conformational rearrangement upon 2A binding.

To investigate this further, we conducted smFRET experiments on freely diffusing molecules to explore the conformations sampled by the frameshifting RNA. We designed two FRET reporter RNAs with the Cy5/Cy3 FRET pair: an 'end reporter' labelled with 5' Cy5 and 3' Cy3 fluorophores (**Figure 3C** and **Figure S2H**), and an 'internal reporter' labelled with 5' Cy5 and an internal Cy3 fluorophore attached to U31 within L2 of the pseudoknot (**Figure 3D** and **Figure S2I**). These were designed to minimally impact RNA folding, avoid dye quenching by adjacent guanine bases, and generate distinct FRET efficiencies in both stem-loop and pseudoknot conformations. The use of alternating-laser excitation (ALEX) confocal imaging⁵⁹ allowed for discrimination between singly- and doubly-labelled FRET-active forms, enabling us to exclusively conduct our downstream analyses on *bona fide* FRET events.

The end reporter RNA exhibited one major high-FRET species (**Figure 3E**, $E_2 = 0.96 \pm 0.01$, ~ 31 Å interdyer distance), consistent with a stem-loop conformation. A second minor FRET population ($E_1 = 0.86 \pm 0.03$, ~ 39 Å interdyer distance) could indicate mobility of the 5' single-stranded region in some molecules, consistent with the SAXS data for the RNA alone (**Figures S2C, S2D**). In the presence of 2A, we observe a single mid-FRET population ($E = 0.63 \pm 0.01$) accompanied by a broadening of the FRET distribution (**Table S3**), consistent with the ends moving further apart (~ 49 Å) and becoming more mobile. This agrees with our crystal structure, in which neither the 5' nor 3' ends of the RNA are stabilised by direct contacts with protein. We do not observe this mid-FRET population in the RNA-only measurements, suggesting that the RNA does not sample the pseudoknot conformation in the absence of 2A.

FRET measurements of the internal reporter RNA revealed a single, mid-FRET population ($E = 0.58 \pm 0.01$) (**Figure 3F**), consistent with a stem-loop conformation (~ 50 Å interdyer distance). Upon addition of 2A, we observe a shift to a major high-FRET population ($E = 0.83 \pm 0.01$) as the dyes move closer together (~ 41 Å) during pseudoknot formation. The FRET distribution also narrows, implying that complex formation stabilises the position of the internal dye. A minor second FRET population is also induced by 2A binding ($E = 0.67 \pm 0.02$, ~ 47 Å). To demonstrate that these changes in FRET efficiency are due to specific interactions with 2A protein and not artefacts of molecular crowding, we performed control experiments with inactive 2A_{R85A/R87A} protein harbouring arginine loop mutations. For both reporters, addition of 2A_{R85A/R87A} results in FRET populations identical to RNA-only experiments (**Figures 3E, 3F**). We also confirmed that ionic strength or identity (Na^+ , K^+ , Mg^{2+}) have no effect on the conformation of the RNA (**Figure S3**): in the absence of 2A protein, the stem-loop conformation persists, irrespective of salt concentration.

Given that conformational plasticity of stimulatory elements has been proposed to regulate frameshifting^{38–41}, we implemented burst variance analysis (BVA)⁶⁰ to further examine dynamics. Briefly, BVA compares the standard deviation of FRET from individual molecules over the timescale of diffusion to the theoretical shot-noise distribution. Our results with both reporters confirm that the RNA does not spontaneously sample multiple stable conformations (**Figure S4**). The end reporter RNA is generally more dynamic than the internal reporter RNA, as indicated by the standard deviation being greater than expected from shot-noise. However, this is likely a feature of dye placement rather than a physiologically relevant characteristic. Overall, our SAXS and FRET analyses support a model in which the stimulatory RNA behaves like a binary switch between two major states: It exists as a stem-loop in isolation, but flips into a pseudoknot in the presence of 2A.

Pseudoknot formation depends on stem 1 sequence, and stem 2 length

To dissect the molecular rules that govern pseudoknot formation, we created a series of mutant RNAs and assessed their binding to 2A by electrophoretic mobility shift assay (EMSA) and microscale thermophoresis (MST) (**Figures 4A, 4B**). We first tested the importance of the specific arginine-guanine interactions observed in the structure (**Figure 2**). Mutations were designed to preserve S1 base pairing and the unusual helical geometry (including the flipped-out base) but remove the possibility of direct arginine-guanine contacts (**Figure 4C**). These included a complete stem swap, and substitution of G-C pairs with A-U and U-A pairs. As assessed by EMSA and MST, none of these mutants could bind to 2A (**Figure 4C**), demonstrating strong specificity for guanine. We also deleted the flipped-out base (Δ U19) allowing S1 to adopt a more regular conformation. However, this did not detectably bind to 2A, highlighting the importance of irregular S1 geometry.

Similarly, we tested the limits of S2 on pseudoknot formation. S2 mutants included a stem-swap, insertion of an additional base pair, and deletion of 1 or 2 base pairs (**Figure 4D**). The stem-swap RNA could bind 2A with similar affinity to wild-type RNA (**Figure 4D**, $K_D \sim 313$ nM, assessed by EMSA). Addition or removal of 1 bp from S2 (creating a 6 or 8 bp S2) had modest effects (~ 313 – 625 nM K_D , assessed by EMSA) but shortening S2 to 5 bp prevented 2A binding, suggesting that length rather than sequence of S2 is important, with 7 bp optimal but 6–8 bp tolerated. The MST data showed a similar trend, but reported lower affinities than EMSA (**Figure 4D**), possibly due to sample aggregation or interactions with capillary surfaces at micromolar concentrations.

Finally, we tested the importance of 2A residues other than R85 and R87 at the pseudoknot interface (**Figure 2** and **Figure S5A**). H26A and R7A mutations had a mild impact on RNA binding (**Figure S5B, S5C**). K24A and W89A exhibited stronger effects, consistent with the specific RNA contacts at these positions, although smFRET experiments showed that each of the four mutations could still trigger pseudoknot formation (**Figure S5D**). Therefore, the only absolute protein requirement is the conserved 2A arginine loop.

Disruption of stem 1 formation inhibits ribosomal frameshifting *in vitro* and in cells

To confirm the mechanistic relevance of our structural observations, we next tested the effects of these mutations on -1 PRF during translation of reporter mRNAs, where ribosomes will encounter the pseudoknot-2A complex during elongation (**Figure 5**). All mutations to S1 and S2 were made in the context of nucleotides 4137–4307 of the TMEV genome, comprising the stimulatory element, slippery sequence and ~ 35 codons upstream to capture any contributions from the nascent peptide. First, we carried out translation *in vitro* using rabbit reticulocyte lysate (**Figures 5A, 5B**). As expected, PRF was 2A-dependent and highly efficient for the wild-type sequence ($91.7 \pm 2.7\%$). All S1 mutants reduced PRF efficiency to background levels ($<2\%$) except Δ U19, which retained some activity ($13.8 \pm 1.7\%$), albeit still 2A-dependent. In contrast, S2 mutants had minor effects on frameshifting. The stem-swap ($89.9 \pm 1.7\%$) and shortening by -1 bp ($87.7 \pm 5.8\%$) led to insignificant changes, whereas lengthening by +1 bp ($48.9 \pm 21.3\%$) or shortening by -2 bp ($36.9 \pm 4.0\%$) caused larger decreases, consistent with trends in 2A-binding affinity (**Figure 4**). We also tested the effects of 2A protein mutations H26A, R7A, K24A and W89A on PRF (**Figure S5E-G**). Only minor effects were observed (59.1 – 82.1% PRF efficiency), consistent with the ability of these variants to bind to wild-type RNA (**Figure S5B-D**).

Next, we assessed the effects of S1 and S2 mutations on frameshifting in HEK293T and BHK-21 cells. We co-transfected cells with a dual-luciferase reporter plasmid harbouring the TMEV

frameshift site, and an expression plasmid for either wild-type 2A or inactive 2A_{R85A/R87A} (Figure 5C). In these experiments, firefly luciferase (F-Luc) is produced by a -1 frameshift, and frameshifting efficiency is calculated relative to an in-frame control (ifc) where R-luc and F-Luc are in the same reading frame (STAR Methods). Notably, frameshifting efficiency on wild-type RNA was lower than *in vitro* (23.6±4.7% in BHK-21, Figure 5D and 17.8±0.9% in HEK293T, Figure 5E), possibly due to our inability to control cytoplasmic 2A concentration. Nevertheless, assays with both cell lines yielded similar results (Figure 5D, 5E). All S1 mutations were strongly inhibitory to -1 PRF (<1.6%) and the ΔU19 mutation had no residual activity in cells, in contrast to *in vitro* experiments. S2 mutants also displayed similar trends: the stem-swap or shortening by -1 bp were mildly inhibitory to -1 PRF, whereas lengthening by +1 bp or shortening by -2 bp were strongly inhibitory. Overall, our translation assays *in vitro* and in cells support the critical importance of 2A-dependent S1 formation as the primary determinant of ribosomal frameshifting.

The first base pair of stem 1 exhibits intrinsic plasticity

In many PRF stimulatory pseudoknots, S1 stability is essential for frameshifting — especially of the first base pairs to be encountered by the ribosome^{61–63}. Our observation that only the second and third base pairs of S1 are stabilised by the arginine loop (Figure 2) was therefore unexpected. We wondered whether single-nucleotide mutations in S1 would prevent rearrangement into the pseudoknot, and designed a series of smFRET reporters to evaluate the relative importance of base pairing at all three positions (Figure 6A). Instead of mutating the guanines that interact with R85 and R87, we targeted the cytosines on the opposite strand (C38G, C37G and C36G), with the triple mutant (CCC→GGG) as a control for completely preventing pseudoknot formation. All individual S1 mutations abolished 2A binding by EMSA and MST (Figure 6B, 6C) and none could adopt the characteristic mid-FRET state ($E = 0.63$, Figure 3E) corresponding to pseudoknot formation (Figure 6D), emphasising the importance of base pairing at all three positions. However, addition of 2A to the C38G mutant produced a minor population of molecules with a pseudoknot-like conformation ($E_1 = 0.76$). This was also present, to a lesser extent, in the C36G dataset ($E_1 = 0.78$). These could not be detected by ensemble methods, consistent with them being unstable, short-lived intermediates. Nevertheless, this suggests some intrinsic plasticity at the first and last position of S1.

The 2A-pseudoknot interaction is likely universal in cardioviruses

The S1-forming GGxG and CCC motifs are highly conserved in cardiovirus genomes⁴⁹. Furthermore, although 2A protein sequences are divergent within the genus, a pair of arginines (equivalent to TMEV R85 and R87) are conserved across *Cardiovirus theileri*, *Cardiovirus rueckerti*, *Cardiovirus saffoldi*, *Cardiovirus dhusarah* and *Cardiovirus rudhira*⁵². To investigate this further, we turned our attention to EMCV (*Cardiovirus rueckerti*), for which the structure of 2A protein is known⁵³ but the minimal RNA stimulatory element is 10 nt larger than that of TMEV (Figure S6A). Structural alignment of the EMCV 2A *apo* structure to the 2A-pseudoknot complex reported here showed that the RNA binding surface is likely compatible with S1 formation (Figure S6B). EMCV 2A arginine loop residues R97 and R95 are juxtaposed with the second and third base pairs of S1 respectively, and the required flexibility for the “G-clamp” is also present: a 5.9 Å movement at glycine 33 (equivalent to TMEV 2A glycine 27) is observed between different chains in the asymmetric unit related by non-crystallographic symmetry.

To test whether the EMCV stimulatory element forms an equivalent pseudoknot upon 2A binding, we designed a doubly-labelled Cy5/Cy3 ‘end reporter’ RNA (**Figure S6C**) and conducted smFRET measurements in the presence and absence of EMCV 2A protein (**Figure S6D**). We observe a major high FRET state in the absence of 2A ($E_2 = 0.98 \pm 0.01$) which decreases ($E = 0.60 \pm 0.02$) in the presence of 2A. To test whether this change is due to pseudoknot formation, we conducted experiments with a C47G mutant RNA (equivalent to TMEV C36G) to disrupt S1. This rendered the EMCV RNA unresponsive to 2A protein (**Figure S6E**).

By manually enforcing S1 base pairs, we used SimRNA⁶⁴ to predict the structure of the EMCV pseudoknot (**Figure S6F**). As expected, the model has a side-by-side arrangement of S1 (3 bp) and S2 (7 bp), and a flipped-out A19 base in S1. The additional 10 nucleotides in EMCV correspond to an extended L2 sequence, predicted to form a tetraloop-capped 4 bp stem that stacks against S2. The absence of L1 in EMCV, but the introduction of an L3 (U48) alters the angle between S1 and S2, making the helical axes almost parallel. However, manual docking of this predicted RNA onto the *apo* structure of EMCV 2A illustrates that the principles of RNA recognition are identical (**Figure S6G**). Overall, the pseudoknot architecture and its binary switching mechanism are likely conserved across all cardioviruses.

Discussion

An unusual RNA pseudoknot architecture

Here we describe a previously unknown class of RNA pseudoknot, the formation of which is strictly conditional upon specific interactions with 2A protein. The inability of AlphaFold3 to predict this complex underscores an urgent need for more experimental structures of diverse RNAs and RNA-protein complexes, to better inform deep learning-based approaches^{42,65}.

In the protein-bound state, the TMEV stimulatory element topologically resembles an H-type pseudoknot, with a 3 bp S1 connected to a 7 bp S2. However, our structure highlights major differences to structurally characterised PRF pseudoknots. Many of these possess two coaxially stacked stems (S1 and S2), exemplified by beet western yellow virus (BWYV; *Potterivirus BWYV*)^{26,66}, pea enation mosaic virus (PEMV; *Enamovirus PEMV*)⁶⁷, potato leaf roll virus (PLRV; *Potterivirus PLRV*)²⁸, sugarcane yellow leaf virus (ScYLV; *Potterivirus SCYLV*)³¹ and SARS-CoV-2 (*Betacoronavirus pandemicum*)^{34,35}. In these structures, folding is nucleated by a stable 5' S1 of at least 5 bp, reinforced by tertiary interactions with nucleotides in loops (reviewed in⁵⁸). In the TMEV RNA, the absence of L3 makes coaxial stacking of S1/S2 impossible: instead they exist in a side-by-side arrangement, with a $\sim 55^\circ$ rotation between helical axes, and L2 adenine nucleotides contribute to tertiary interactions with S1.

Why does this pseudoknot depend on 2A protein? In RNA-only experiments, the molecule exists exclusively as a stem-loop (S2), and S1 does not form, even though intuitively the pseudoknot (10 bp, 27 Watson-Crick hydrogen bonds) appears to be more stable than the stem-loop (7 bp, 18 Watson-Crick hydrogen bonds). Our structure suggests that the answer to this question is related to the sub-optimal geometry of S1. When RNA pseudoknots were first described, three base pairs were proposed as a theoretical minimum for S1 stability⁶⁸, and our discovery of a flipped-out base within the 3 bp S1 was unexpected. The energetic penalty associated with displacement and solvation of U19 likely explains the dependency of S1 formation on favourable arginine-guanine interactions. In this way, 2A can be thought of as a catalyst or chaperone, reducing the energy barrier for the stem-loop to pseudoknot transition. Once formed, the pseudoknot is further stabilised by interactions with 2A (8 protein-RNA

hydrogen bonds, 634 Å² interface) consistent with the contribution of ΔH (-10.7±0.13 kcal/mol) to the overall ΔG of 2A binding (-9.8 kcal/mol)⁵². This explains the ability of the complex to function as a blockade during ribosome elongation.

High-efficiency frameshifting in cardioviruses

Why is the TMEV frameshift efficiency so high? Numerous biochemical and virological studies have identified the stability of S1 — in particular, the first base pairs of S1 — as a major determinant of frameshifting^{61–63}. A striking feature of the TMEV 2A-RNA complex is that the arginine loop exclusively stabilises the second and third positions of S1, not the first. Furthermore, our smFRET data suggests that transient 2A-RNA interactions may still occur in the absence of the first base pair. We hypothesise that as the ribosome approaches S1, reversible breaking and reforming of the first base pair may occur before 2A dissociates, thereby creating tension on the mRNA that favours a -1 frameshift. This could contribute to the unusually high PRF efficiency (~85%) in cardioviruses that cannot be explained by thermodynamic equilibrium models^{22,23}. Additional evidence for mRNA tension in this system comes from the observation that reducing the spacer length by one nucleotide biases the ribosome towards -2 frameshifting⁵¹, also observed at the HIV-1 *gag/pol* site¹⁹.

Our structure also resolves a mystery regarding the unusually long spacer between the PRF slippery sequence and downstream stimulatory element. This was initially reported to be 14 nt, based on bioinformatic identification of S2 and experiments demonstrating its importance for PRF^{48,50}. However, our discovery of S1 in the 2A-RNA complex reduces the spacer to 9 nt between the end of the GGUUUUU slippery sequence the start of the GGxG motif comprising the 5' strand of S1, comparable to other PRF signals². Interestingly, the first contact between 2A and the PRF stimulatory RNA is two nucleotides earlier than S1, with A15 stacking against W89. Although this is not required for 2A binding (**Figure S5**) it may nevertheless contribute to slowing down elongation as the ribosome approaches the first base pair of S1¹⁴.

A protein-dependent riboswitch?

In this system, the RNA behaves as a binary switch, converting from stem-loop to pseudoknot in the presence of 2A. In contrast, the 2A protein undergoes minimal conformational change during complex formation (C_α backbone RMSD of 0.39 Å). We therefore propose this element represents a protein-dependent riboswitch. Generally, riboswitches modulate gene expression through interconversion between structurally distinct conformations in response to a ligand (reviewed in^{58,69,70}). To date, riboswitches have been described for a wide variety of ligands, including nucleotide precursors (pre-queuosine₁, preQ₁⁷¹), metabolites (coenzyme B12⁷², S-adenosylmethionine, SAM^{73,74}), amino acids (lysine⁷⁵, glycine⁷⁶, glutamine⁷⁷) and even tRNAs (T-box⁷⁸). In our case, the 2A protein itself acts as the primary ligand. Although this is unusual, ribosomal proteins have been shown to modulate the stability of adenine-sensing riboswitches⁷⁹. Another conceptual difference here is that the TMEV PRF RNA exerts its regulatory effect at the level of translation elongation, rather than translation initiation or transcription termination.

A feature essential to riboswitches is the stabilisation of the ligand-sensing 'aptamer' domain, typically S2, by the cognate ligand. This is common in bacterial riboswitches that regulate translation initiation, as the Shine Dalgarno sequences are occluded within the stabilised S2. In our case the 'aptamer' would be the GGxG motif comprising S1, and the ligand would be the arginine loop of 2A. The selectivity for arginine is demonstrated by the complete inability

of the 2A_{R85A/R87A} protein to induce switching (**Figure 3**) and no evidence for spontaneous sampling of the pseudoknot conformation in the absence of protein. This contrasts with the preQ₁ riboswitch, which transiently samples an intrinsically unstable pseudoknot-like fold in the absence of preQ₁ ligands⁸⁰.

In conclusion, this work defines the molecular basis for the highest efficiency viral frameshifting event in nature. Reframing the idea of a PRF stimulatory element as a riboswitch has vast implications. We speculate that other canonical PRF elements — especially those reported to exhibit multiple conformations — might also be regulated by cellular proteins, viral proteins or metabolites. Indeed, ribosome profiling of cells infected with mouse hepatitis virus (MHV; *Betacoronavirus muris*⁸¹), SARS-CoV-2^{82,83} and PRRSV⁸⁴ has revealed variations in frameshifting over time at PRF signals previously thought to function at a fixed efficiency. Understanding this hidden layer of regulation will be essential to realise the potential of artificial intelligence in designing ligands to manipulate these elements^{85,86}.

Limitations of the Study

While the SAXS and FRET data support the presence of a single RNA conformation in the absence of 2A, the possibility of transient pseudoknot sampling under different conditions (ionic strength, temperature, molecular crowding, or in the presence of the ribosome) cannot be excluded. Differences in gaussian peak widths observed in our smFRET measurements could reflect either a broader ensemble of RNA conformations, enhanced flexibility in the 5' or 3' ends and/or increased dynamic interconversion between states occurring on timescales faster than the measurement resolution. Our current data do not allow us to unambiguously distinguish between static and dynamic contributions.

Resource Availability

LEAD CONTACT

- Requests for further information and resources should be directed to and will be fulfilled by the lead contact, Chris H. Hill (chris.hill@york.ac.uk).

MATERIALS AVAILABILITY

- Plasmid maps generated in this study have been deposited to Mendeley Data **10.17632/78kj389b6v.1**.

DATA AND CODE AVAILABILITY

- Crystallographic data and refined atomic models for the 2A-RNA complex have been deposited to the Worldwide Protein Data Bank (wwPDB) under accession code 9RVP. Small-angle X-ray scattering data has been deposited to the Small Angle Scattering Biological Data Bank (SASBDB) under accession codes SASDWQ5 (2A protein only), SASDWR5 (RNA only) and SASDWS5 (2A-RNA complex). Original gel images, smFRET raw hdf5 data files and Jupyter Notebooks used for smFRET data analysis and have been deposited at Mendeley Data **10.17632/78kj389b6v.1**. These data are publicly available as of the date of publication.
- This paper does not report original code.

- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Acknowledgments

We acknowledge the YSBL X-ray crystallography facility supported by BBSRC (BB/T017805/1), Diamond Light Source for access to beamline I04 under proposal mx-32736, and EMBL for access to the bioSAXS beamline P12 under proposal SAXS-1125. We thank Johan Turkenburg and Sam Hart for technical assistance at the YSBL X-ray crystallography facility. Remote synchrotron access was supported in part by the EU FP7 infrastructure grant BIOSTRUCT-X (Contract No. 283570). We thank the Molecular Interactions team (Andrew Leech, Katy Cornish and Alex Payne-Dwyer) and the Protein Production team (Jared Cartwright and Matt Tyreman) for technical support within the York Bioscience Technology Facility. We also thank Ben Butt for helpful discussions. J.K.B. is supported by a Medical Research Council DiMeN PhD studentship (MR/W006944/1). C.M.J. acknowledges the European Union's Horizon 2020 research and innovation programme "iNEXT Discovery" grant agreement no. 871037. M.A.S.A. was supported by BBSRC grant to TDC (BB/T008032/). T.D.C. is supported by BBSRC (BB/W00061X/1). M.C.L. is supported by BBSRC (BB/W000555/1) and EPSRC (EP/Y000501/1). S.D.Q. is supported by Alzheimer's Research UK (RF2019A-001) and EPSRC (EP/X525856/1, EP/V034030/1). S.P.G. was supported by an EPSRC studentship (EP/T518025/1). H.C.Y.K. was funded by an Exciting Instruments (Sheffield, UK) studentship administered by Generation Research (University of York, UK). S.C.G. was funded by a Sir Henry Dale fellowship (098406/Z/12/B) from the Wellcome Trust and the Royal Society. We acknowledge BBSRC funding to I.B. and C.H.H. (BB/V000306/1). T.C.P. and C.H.H. are supported by a Sir Henry Dale Fellowship (221818/Z/20/Z) from the Wellcome Trust and the Royal Society (to C.H.H.). C.H.H. is supported by the Lister Institute of Preventive Medicine. For the purpose of open access, the authors have applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

Author Contributions

Conceptualization, J.K.B., C.H.H. Data Curation, J.K.B., C.M.J., C.H.H. Formal Analysis, J.K.B., C.M.J., T.C.P., C.H.H. Funding Acquisition, S.C.G., C.H.H. Investigation, J.K.B., C.M.J., T.C.P., H.C.Y.K., S.P.G., I.B., C.H.H. Methodology, J.K.B., C.M.J., M.A.S.A., Project Administration, S.D.Q., C.H.H. Resources, J.K.B., C.M.J., I.B., C.H.H. Software, M.A.S.A., T.D.C., Supervision, M.A.S.A., J.A.L.H., M.C.L., S.D.Q., C.H.H. Validation, J.K.B., C.M.J., C.H.H. Visualization, J.K.B., C.H.H. Writing – original draft, J.K.B., C.H.H., Writing – review & editing, J.K.B., T.D.C., S.C.G., I.B., M.C.L., S.D.Q., C.H.H.

Declaration of Intrests

T.D.C. and M.A.S.A. are employees and shareholders in Exciting Instruments Ltd.

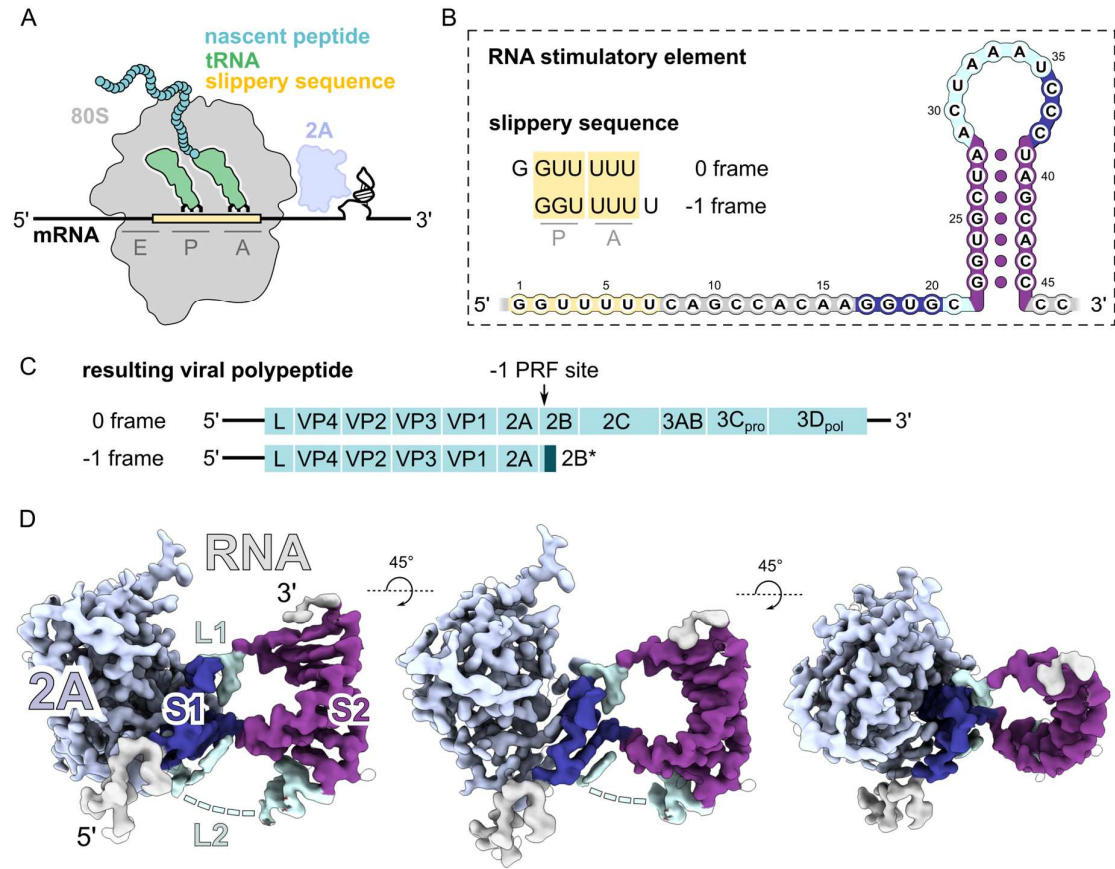


Figure 1. The PRF stimulatory element from TMEV is a protein-pseudoknot complex

A) Diagram of -1 PRF in the cardioviruses. Frameshifting is conditional on the virally encoded 2A protein. The 2A protein (light blue) binds to the structured RNA stimulatory element (black, predicted stem-loop), stalling the elongating ribosome whilst the slippery sequence (yellow) is positioned in the decoding centre. tRNAs (green) may uncouple and recouple in the -1 frame, prior to continuation of the ribosome.

B) Details of the TMEV frameshift site. Numbering starts from slippery sequence (yellow). The stimulatory element is predicted to form a 7 bp stem-loop (purple). Conserved GGxG and CCC motifs (dark blue) and divergent interconnecting sequences (cyan) are highlighted.

C) The effect of -1 PRF on viral gene expression. Leader protein (L) and capsid proteins (VP4, VP2, VP3, VP1) are encoded 5' to the -1 PRF site (indicated by black arrow) whereas the replicative proteins (2B, 2C, 3AB, 3C_{pro} – protease, 3D_{pol} – RNA polymerase) are encoded 3' to the -1 PRF site. Frameshifting results in the truncation of the viral polypeptide, creating the 2B* product, and downregulating expression of the replicative proteins. Up to 85% of elongating ribosomes are diverted into the -1 frame.

D) 1.9 Å X-ray crystal structure of the 2A-RNA complex reveals that the RNA adopts a pseudoknot fold, not a stem-loop, when bound to 2A. Electron density (2F_o-F_c) within 1.2 Å of the refined models is shown in three views. 5' and 3' nucleotides not involved in base pairing are shown in grey, stem 1 (S1) in blue, stem 2 (S2) in purple, and loops (L1, L2) in cyan, corresponding to the colouring of the sequence in B). Three disordered nucleotides in L2 are shown as a dashed line.

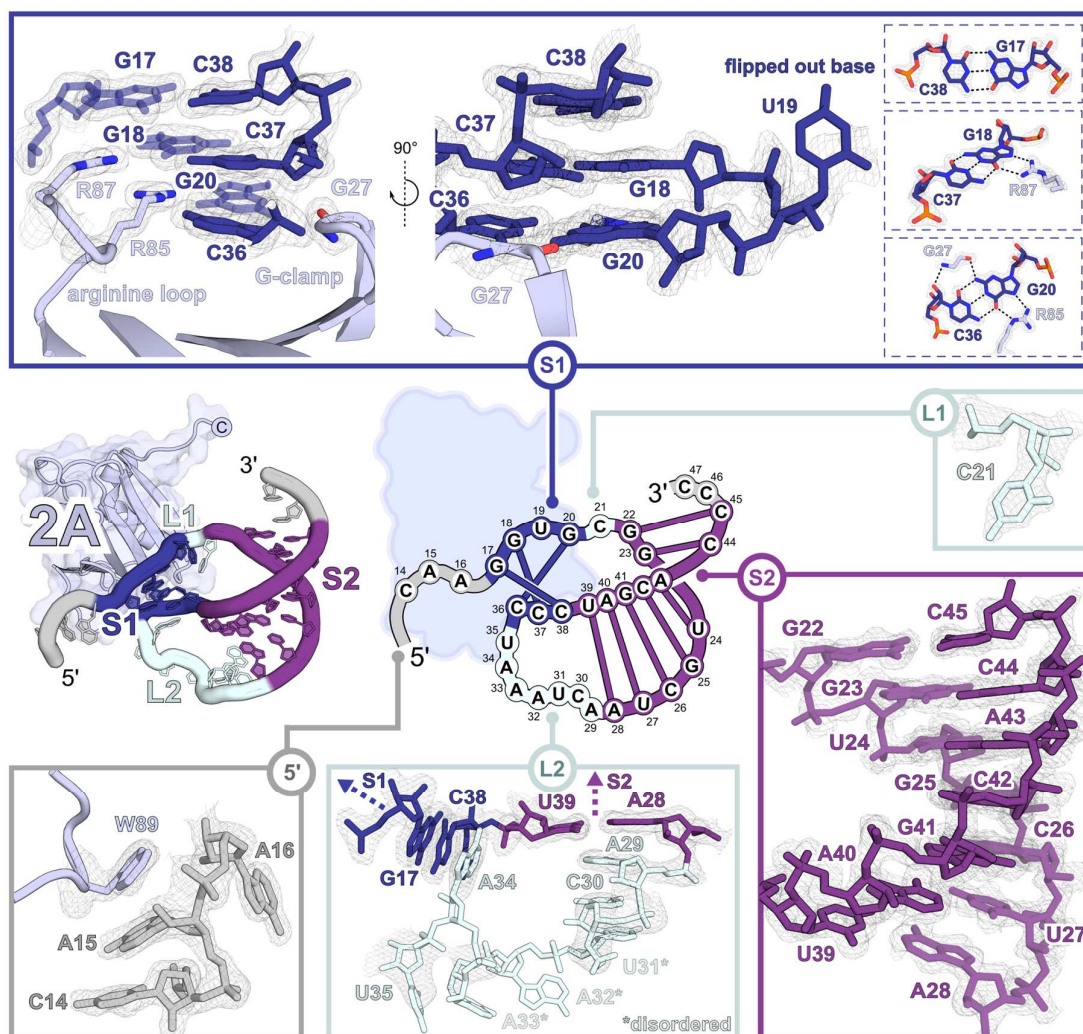


Figure 2. Structural features of the RNA pseudoknot in complex with 2A

Close-up views illustrating key regions of the RNA pseudoknot (also see [movie S1](#)). The structure is shown as sticks, modelled into $2F_o - F_c$ electron density (grey mesh) contoured at $0.64 \text{ e}^-/\text{\AA}^3$ (2.0 RMSD). (Centre) Topological diagram of the pseudoknot with base identity indicated. S1 consists of three G-C Watson-Crick base pairs comprised of the conserved GGxG and CCC triplet motifs, supported by the arginine loop (R85 and R87) and 'G-clamp' (G27) of 2A. The flipped base (U19) allows the formation of S1. Details of hydrogen bonding in S1 are shown. L1 is a single nucleotide (C21). S2 is a seven Watson-Crick base paired A-form RNA helix. L2 is mostly disordered. The 5' end of the RNA stacks against W89.

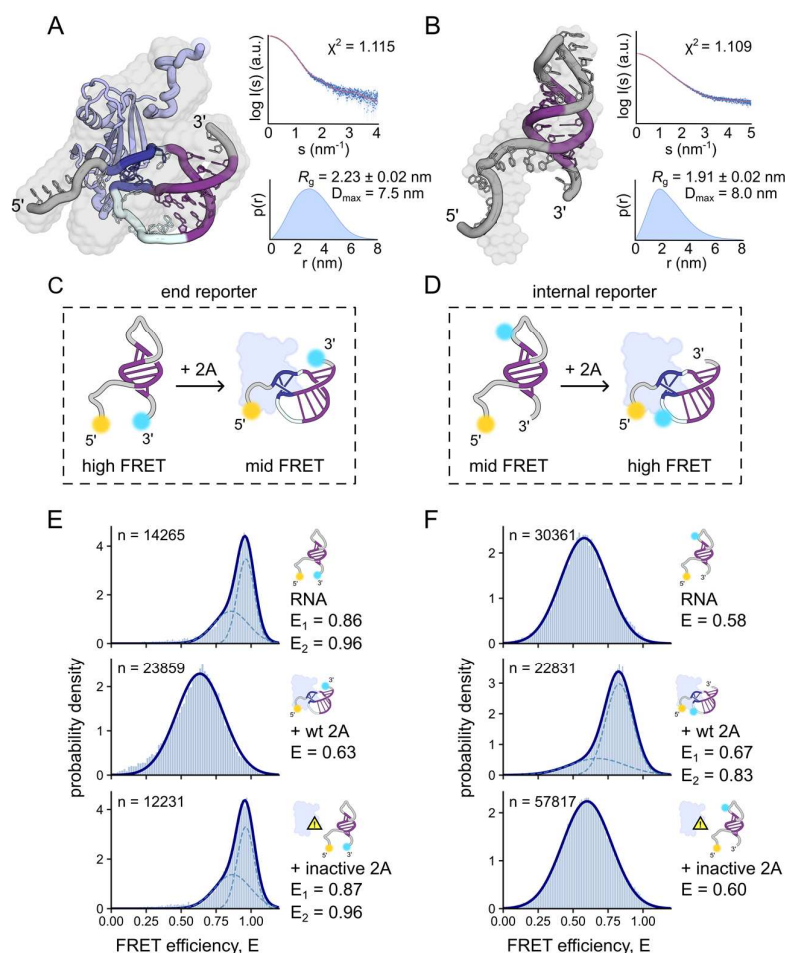


Figure 3. The RNA stimulatory element switches conformation from stem-loop to pseudoknot upon 2A binding

A,B) SEC-SAXS envelope with plots showing alignment between experimentally collected SAXS data (blue circles) and calculated data from structural models (red lines), and the pair-distance distribution $P(r)$ allowing the determination of $R_g \pm \text{SD}$ and D_{max} . A) 2A-RNA complex with X-ray crystal structure fitted to the data. B) RNA in isolation, supporting a stem-loop model.

C,D) Diagrams showing smFRET RNA reporter construct design using the Cy3 (cyan) and Cy5 (yellow) FRET pair.

E,F) Corrected FRET efficiency (E) histograms combined from three biological repeats of the E) end reporter and F) internal reporter wild-type RNAs. Showing data for (*upper*) RNA in isolation, (*middle*) RNA in the presence of $0.5 \mu\text{M}$ wild-type 2A and (*lower*) RNA with $0.5 \mu\text{M}$ inactive 2A_{R85A/R87A}. Corrected average FRET efficiencies (E) and number of molecules (n) are reported in each histogram. Gaussian fits are shown as dashed curves for two Gaussians and a solid blue line for a single Gaussian. $R^2 = 0.99$ for all gaussian fits.

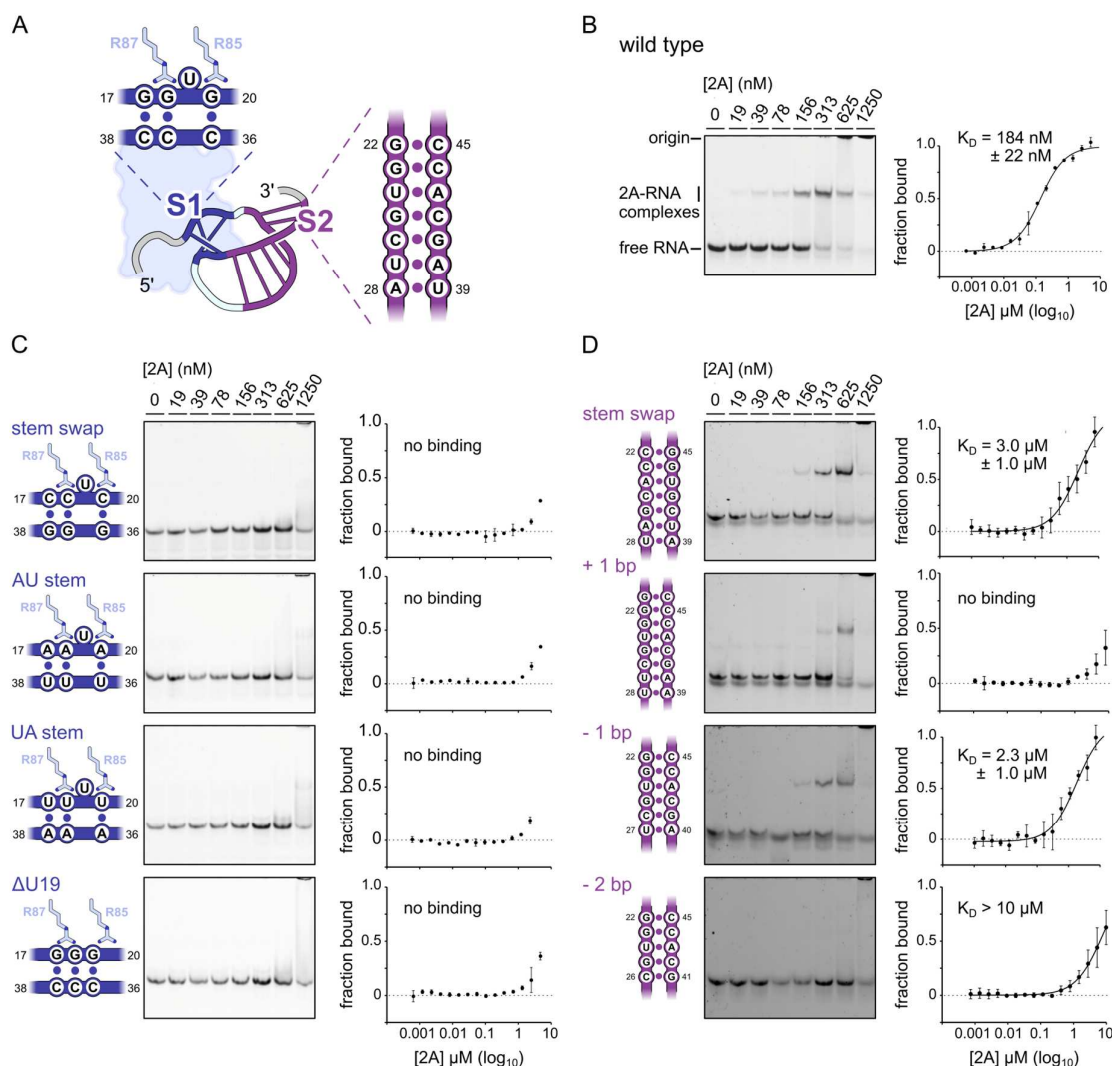


Figure 4. Pseudoknot formation depends on stem 1 sequence, and stem 2 length

A) Diagrams showing the location of mutated nucleotides in stem 1 (S1) and stem 2 (S2).

B) (Left) EMSA analysis of 2A protein binding to wild-type stimulatory RNA. (Right) MST binding curves and approximate K_D values using the same RNA, assuming 1:1 binding stoichiometry. Experiments were independently performed twice, with error bars indicating $\pm$ SD.

C) (Left) Details of each S1 mutation. (Middle) EMSA analysis of the effects of S1 mutations on 2A binding, as in B). (Right) MST binding curves and approximate K_D values using the same S1 mutant RNAs, as in B).

D) (Left) Details of each S2 mutation. (Middle) EMSA analysis of the effects of S2 mutants on 2A binding, as in B). (Right) MST binding curves and approximate K_D values using the same S2 mutant RNAs, as in B).

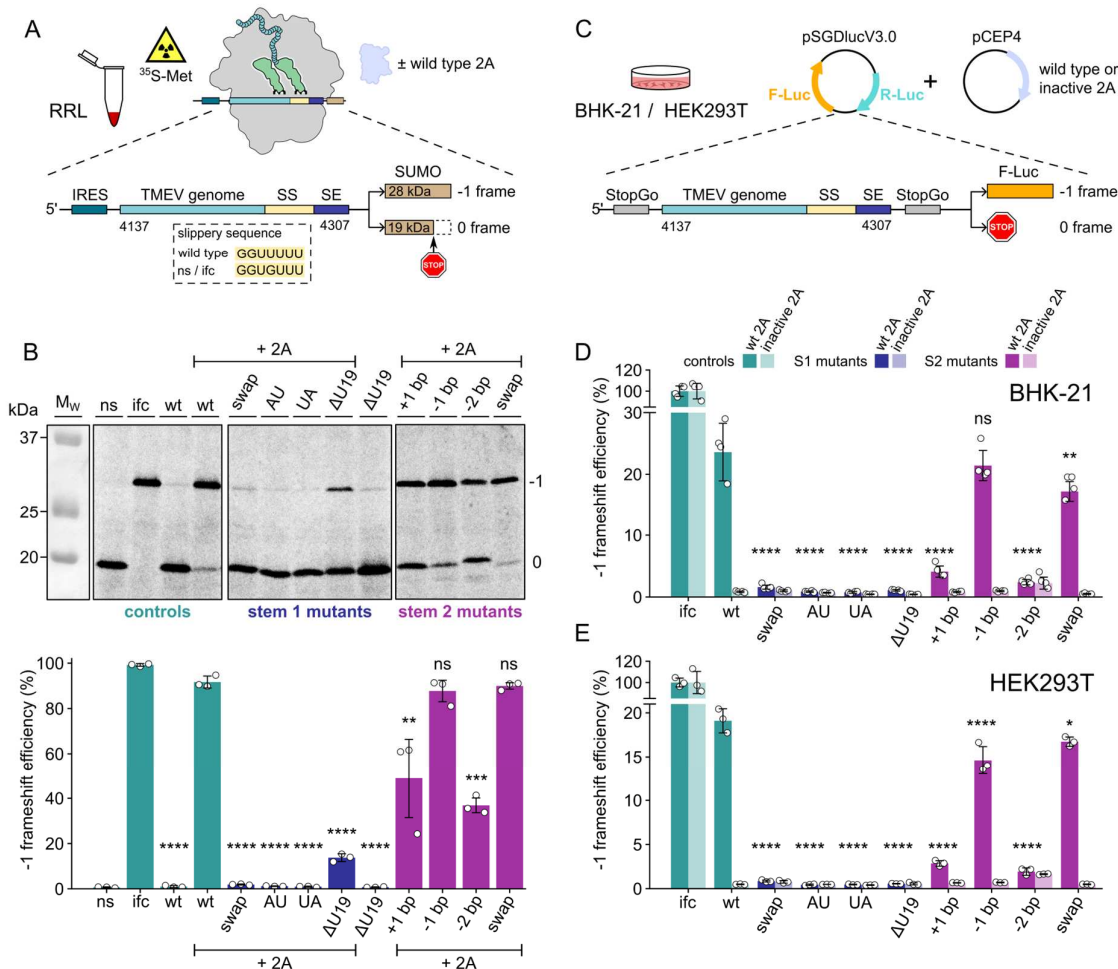


Figure 5. The stem 1 sequence is essential for -1 PRF during translation *in vitro* and in cells

A) Diagram of *in vitro* translation assay to measure frameshifting efficiency. Translation reactions were carried out in rabbit reticulocyte lysate (RRL) in the presence of ^{35}S -Met. Reporter mRNAs include 35 codons upstream of the frameshift site (light blue), the slippery sequence (SS, yellow) stimulatory element (SE, dark blue) and SUMO coding sequence (brown) in the -1 frame.

B) (Upper) SDS-PAGE visualised by autoradiography, showing effects of stem 1 and 2 mutations on frameshifting efficiency *in vitro*. (Lower) Densitometric quantification of -1 frameshift efficiencies, corrected for methionine content. Data are represented as the mean $\pm$ SD error bars ($n = 3$ biological repeats). ns, not significant, $*p \leq 0.05$, $**p \leq 0.01$, $***p \leq 0.001$, $****p \leq 0.0001$. Control mRNAs: non-slip (ns), in-frame control (ifc) and wild-type (wt) are coloured in teal. Stem 1 mutants are coloured in dark blue. Stem 2 mutants are coloured in dark purple.

C) Diagram of dual luciferase assay used to measure frameshifting efficiency in HEK293T and BHK-21 cells. Cells were co-transfected with two plasmids to transiently express a dual luciferase reporter cassette containing sequences of interest, along with either wild-type or inactive 2A_{R85A/R87A} protein. Elements are colour-coded as in A).

D, E) Calculated -1 frameshift efficiencies from luminometry measurements in D) BHK-21 cells 36 hpt and E) HEK293T cells 24 hpt. Data are represented as the mean $\pm$ SD error bars, normalised to the in-frame control (ifc) ($n = 4$ biological repeats for BHK-21 and $n = 3$ biological repeats for HEK293T). ns, not significant, $*p \leq 0.05$, $**p \leq 0.01$, $***p \leq 0.001$, $****p \leq 0.0001$.

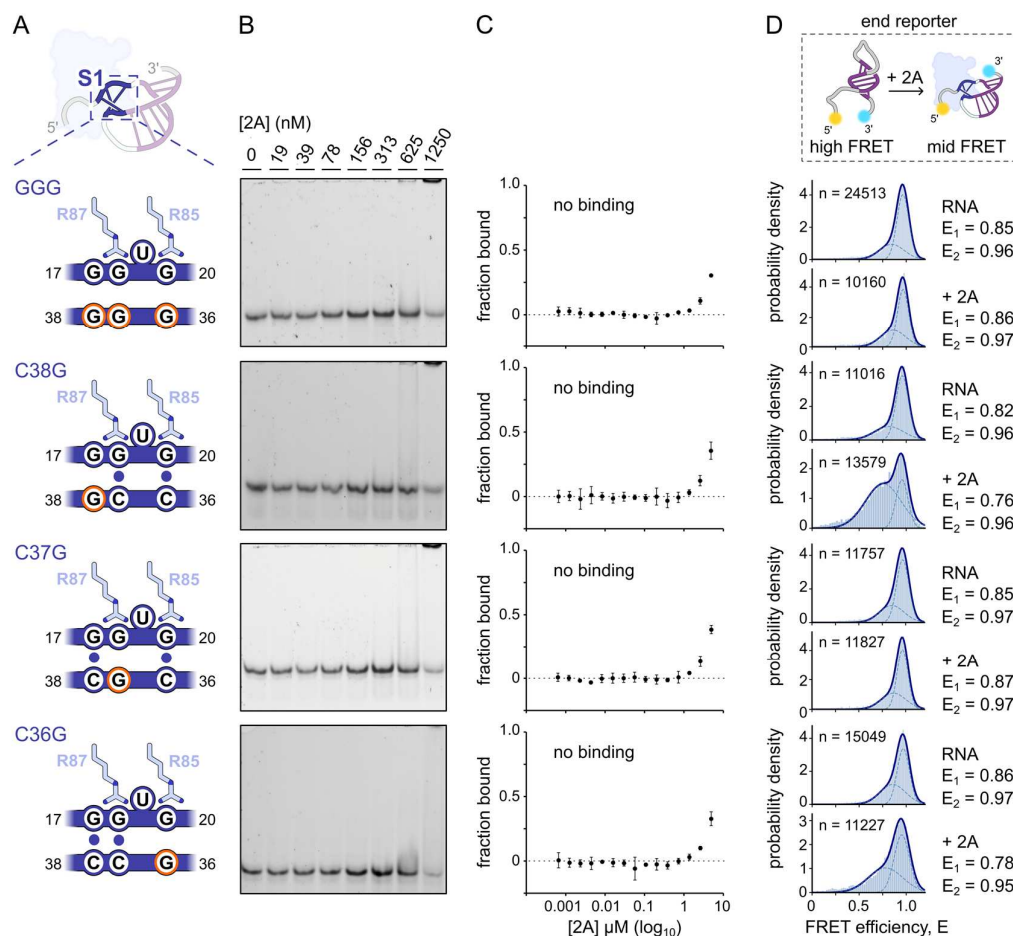


Figure 6. All three base pairs of stem 1 are critical for pseudoknot formation

A) Diagram showing the location of mutated nucleotides in stem 1 (S1) that disrupt base pair formation but not arginine-guanine interactions.

B) EMSA analysis of the effects of disrupting individual S1 base pairs on 2A binding.

C) MST binding curves and approximate K_D values using the same S1 mutant RNAs, assuming 1:1 binding stoichiometry. Experiments were independently performed twice, with error bars indicating $\pm$ SD.

D) Corrected FRET efficiency (E) histograms of S1 mutant end reporter RNAs, combined from three biological repeats. Showing data for (upper) RNA in isolation, (lower) RNA in the presence of 0.5 μM wild-type 2A. Corrected average FRET efficiencies (E) and number of molecules (n) are reported in each histogram. Gaussian fits are shown as dashed curves for two Gaussians and a solid blue line for a single Gaussian. $R^2 = 0.99$ for all Gaussian fits.

601 **Table 1.** Crystallographic data collection and refinement statistics^a

	2A-RNA pseudoknot PDB ID: 9RVP
Data collection	
Wavelength (Å)	0.9537
Space group	C 121
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	104.19, 45.93, 65.72
α, β, γ (°)	90.00, 127.34, 90.00
Resolution (Å)	52.25-1.90 (1.94-1.90)
<i>R</i> _{merge}	0.152 (0.728)
<i>I</i> / σ <i>I</i>	7.1 (3.3)
Completeness (%)	98.7 (98.3)
Multiplicity	6.5 (6.7)
CC _{1/2}	0.974 (0.640)
Refinement	
Resolution (Å)	52.25-1.90 (2.02-1.90)
No. reflections	19410 (3175)
<i>R</i> _{work} / <i>R</i> _{free}	0.1931 / 0.2295
No. atoms	
Protein ^b	2043
RNA ^b	962
Water	249
<i>B</i> -factors	
Protein	28.1
RNA	47.7
Water	37.2
R.m.s. deviations	
Bond lengths (Å)	0.009
Bond angles (°)	0.831
Ramachandran statistics	
Preferred (%)	98.35
Allowed (%)	1.65
Outliers (%)	0.00
All-atom clashscore	1.33
Rotamer Outliers (%)	0.00

602 ^aData were recorded from a single native crystal. Values in parentheses are for highest-resolution
603 shell. ^bIncluding hydrogen atoms.

604 **Table 2.** SAXS data collection and analysis parameters: 2A protein, RNA and 2A-RNA complex

Sample details			
Source Organism	<i>Theiler's murine encephalomyelitis virus (TMEV)</i>		
Scattering particle composition	2A protein – UniProt: P08545 . Recombinantly expressed TMEV genome polyprotein amino acids 923–1055, with additional non-native amino acids at the N- and C-terminus (GPLGS; EFPGRLERPHRD). RNA ^a – 37 polynucleotide single stranded RNA.		
Sample environment			
X-ray sample cell	1 mm pathlength in-vacuum quartz capillary; continuous flow		
Size Exclusion Chromatography SEC-SAXS	2A protein ^b	RNA	2A-RNA complex
Solvent composition	20 mM HEPES pH 7.4, 1 M KCl, 2% v/v glycerol, 1 mM DTT	20 mM HEPES pH 7.4, 400 mM KCl, 2% v/v glycerol, 1 mM DTT	
SEC column temperature	Room temperature, nominally 20°C		
Sample injection concentration	3.96 mg/mL	7.08 mg/mL	7.95 mg/mL
Sample injection volume	50 and 100 µL ^b	50 µL	
SEC column type	GE Superdex 75 Increase 5/150		
SEC flowrate	0.35 mL/min		
SAS data collection			
Data acquisition/reduction software	EMBL bioSAXS SASFLOW pipeline; <i>BECQUEREL</i> beam line control software; <i>CHROMIXS</i> ; Agilent ChemStation software (RRID:SCR_015742).		
Source/instrument description or reference	EMBL P12 bioSAXS beam line (DESY, Hamburg) equipped with a Pilatus 6M detector (Dectris; s-axis calibrated with silver behenate) coupled to an Agilent 1260 Infinity II Bio-Inert HPLC.		
Method for scaling intensities	Relative (instrument) scale, normalized to beam transmission		
	2A protein	RNA	2A-RNA complex
X-ray wavelength (energy)	0.123983 nm (10 keV)		
Measured s-range (s_{min} – s_{max} nm ⁻¹)	0.026–7.41	0.026–7.41	0.026–7.42
Working s-range (s_{min} – s_{max} nm ⁻¹) ^c	0.076–4.00	0.104–5.84	0.104–4.70
Exposure time per data frame (total number of frames)	0.25 s (2880)	0.25 s (2880)	0.25 s (2880)
Number of sample frames used for averaging (SEC elution peak)	143 ^b	57	41
SAS-derived structural parameters			
Methods/Software	<i>PRIMUS</i> (ATSAS 3.0) including <i>GNOM 5</i>		
	2A protein	RNA	2A-RNA complex
Guinier Analysis R_g correlation/stability through SEC-elution peak			
Mean $R_g \pm \sigma$ (nm)	1.89 (0.05)	1.91 (0.02)	2.23 (0.02)
Median R_g (nm)	1.89	1.92	2.23
R_g min-max range (nm)	1.79–2.00	1.88-1.94	2.20–2.26
Guinier Analysis (final scaled and averaged data)			
$I(0) \pm \sigma$ (cm ⁻¹)	1704.7 (3.7)	7530.6 (6.4)	6644.9 (7.2)
$R_g \pm \sigma$ (nm)	1.89 (0.01)	1.92 (0.01)	2.23 (0.01)
$min < sR_g < max$ limit (data point range) ^d	0.14–1.30 (19–239)	0.31–1.23 (49–221)	0.23–1.30 (29–201)

Linear fit assessment (Pearson correlation coefficient ² , R ²)	0.987	0.998	0.998
<i>P(r) analysis</i>	2A protein	RNA	2A-RNA complex
<i>I(0) ± σ (cm⁻¹)</i>	1714 (4.4)	7577 (5.8)	6651 (7.6)
<i>R_g ± σ (nm)</i>	1.94 (0.01)	1.99 (0.01)	2.24 (0.01)
<i>D_{max} (nm)</i>	7.2	8.0	7.5
<i>s-range for fit (nm⁻¹)</i>	0.076–3.96	0.104–5.84	0.104–4.70
<i>P(r) fit assessment c²</i> (CorMap- <i>P</i>)	0.920 (0.288)	1.09 (0.222)	1.05 (0.095)
Scattering particle size			
<i>Methods/Software</i>	ATSAS 3 data analysis modules		
	2A protein	RNA	2A-RNA complex
<i>Volume estimates</i>			
Porod volume, <i>V_p</i> (nm ³) ^e	30	19	46
<i>Molecular weight estimates</i> (kDa)			
MW calculated from sequence	17.4	11.8	29.2 (1:1 complex)
MW from SAXS/ <i>P(r)</i> : (MW range; > 0.9 confidence) ^e	14.8 (14.5–15.8)	13.5 (13.1–14.5)	27.6 (25.3–28.9)
Modelling			
<i>Shape modelling software</i>	MONSA, DAMMIN		
	2A protein	RNA	2A-RNA complex
<i>MONSA</i>			
<i>s-range for fit (nm⁻¹)</i>	0.099–4.00	0.110–4.00	0.105–4.00 (complex) 0.099–4.00 (protein)
Symmetry	P1	P1	P1 (two phase modelling)
# of model reconstructions	5	5	5
χ^2 , CorMap <i>P</i> -values for fit	0.94–0.95 (0.749– 0.288)	1.10–1.11 (0.081)	0.94–0.95 (0.081– 0.748)
		<i>Protein phase</i>	1.05–1.04 (0.0007)
<i>DAMMIN</i>			
<i>s-range for fit (nm⁻¹)</i>	0.099–4.00	0.110–4.00	
Symmetry	P1	P1	
# of model reconstructions	10	10	
Normalized spatial discrepancy (NSD)	0.586	0.746	
χ^2 , CorMap <i>P</i> -values for fit	0.92 (0.745–0.288)	1.10–1.12 (0.218– 0.001)	
<i>Atomistic modelling</i> <i>software</i>	AlphaFold3, SREFLEX, CORAL, CRY SOL		
	2A protein	RNA	2A-RNA complex
<i>Model template(s)</i>	X-ray crystal structure	Alphafold3	X-ray crystal structure
<i>s-range for fit (nm⁻¹)</i>	0.076–3.96	0.082–4.00	0.071–3.50
Symmetry/stoichiometry	P1 (monomer)	P1 (monomer)	P1 (1:1 complex)
# of model reconstructions	5	9	5
χ^2 , CorMap <i>P</i> -values for fit ^f	0.92–0.97 (0.937– 0.156)	1.11–1.20 (0.191– 0.013)	1.11–1.35 (0.021– 0.000)
SASBDB accession codes	SASDWQ5	SASDWR5	SASDWS5

^aRNA sequence: 5' - CCACAAGGUGCGGUGCUAACUAAAUCCCUAGCACCCC - 3'. ^bFor the 2A protein, two consecutive and independent SEC-SAXS injections of different sample volumes were measured. The data, structural parameters etc. for the 2A protein are those derived from the scaled and averaged SAXS data across both SEC-SAXS measurements. ^cThe working s-range was determined using a combination of *AUTORG* ($s_{\min}$)⁸⁷ and *SHANUM* ($s_{\max}$)⁸⁸. ^dThe data point range quoted here is relative to the full data-point/s-range of the measured SAXS profile (not the working s-range). ^ePorod volume and concentration-independent MW estimates from the SAXS data were calculated using the corresponding GNOM.out files as input into the ATSAS 3 tools *DATPOROD* and *DATMW*, respectively. ^fModel-data fit assessment using CRY SOL v. 2.8 (30 spherical harmonics, 256 points, constant enabled; average solvent electron density = 0.3416 e/Å³).

STAR Methods

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
<i>Escherichia coli</i> Stellar cells	Takara	Cat# 636763
<i>Escherichia coli</i> BL21(DE3) pLysS cells	Sigma	Cat# 69451-3
Chemicals, peptides, and recombinant proteins		
2 × TY broth capsules	Melford	Cat# X15640
DMEM	Gibco	Cat# 41966-029
Fetal Bovine Serum	Gibco	Cat# A5256701
Penicillin-Streptomycin (5,000 U/mL)	Gibco	Cat# 15070063
L-Glutamine (200 mM)	Gibco	Cat# 25030149
Isopropyl-β-D-thio-galactopyranoside (IPTG)	Melford	Cat# I56000
Trizma® base	Sigma	Cat# T1503
Sodium chloride	Fisher Chemical	Cat# S/3161/65
Imidazole	Sigma	Cat# 56750
DNase I	Sigma	Cat# D4527
DTT	Melford	Cat# D11000
Potassium chloride	Sigma	Cat# P9541
Glycerol	Honeywell	Cat# 49770-2.5L
HEPES	Fluorochem	Cat# F047861
Ammonium acetate	Sigma	Cat# A1542
Bis-Tris	Sigma	Cat# B9754
Poly(ethylene glycol) 3350	Sigma	Cat# 88276
Magnesium acetate tetrahydrate	Fluorochem	Cat# F449646
Bromophenol Blue sodium salt	Sigma	Cat# B5525
ProtoGel (30%)	National Diagnostics	Cat# EC-890
AccuGel 19:1 (40%)	National Diagnostics	Cat# EC-850
Urea, UltraPure	Melford	Cat# U20200
TEMED	Melford	Cat# T18000
Ammonium persulfate (APS)	Melford	Cat# A20500
Bovine liver tRNA	Sigma	Cat# R4752
UltraPure BSA (50 mg/mL)	Invitrogen	Cat# 10743447
Magnesium chloride, hexahydrate	Millipore	Cat# 442611-M
AzDye647-maleimide	Vector Labs	Cat# FP-1122-1
Cy5-maleimide	GE Healthcare	Cat# PA25031
NdeI	New England Biolabs	Cat# R0111S
dNTP bundle	Jena Bioscience	Cat# NU-1005S
RiboLock RNase inhibitor	Thermo Fisher	Cat# EO0381
L-[S35] Methionine 10 mCi/ml. Translation grade.	Hartmann Analytic	Cat# SRM-01
TransIT®-LT1 transfection reagent	Mirus Bio	Cat# MIR 2304
PEI transfection reagent	MedChemExpress	Cat# HY-K2014
Passive Lysis 5X buffer	Promega	Cat# E1941
Recombinant protein: wild type TMEV 2A	This paper, based on Hill <i>et al.</i> ⁵²	N/A
Recombinant protein: TMEV inactive 2A _{R85A/R87A}	This paper, based on Hill <i>et al.</i> ⁵²	N/A
Recombinant protein: TMEV 2A _{R7A}	This paper	N/A
Recombinant protein: TMEV 2A _{K24A}	This paper	N/A
Recombinant protein: TMEV 2A _{H26A}	This paper	N/A

Recombinant protein: TMEV 2A _{W89A}	This paper	N/A
Recombinant protein: EMCV wild type 2A	This paper, based on the work of Hill <i>et al.</i> ⁵³	N/A
Critical commercial assays		
5' EndTag DNA/RNA Labeling Kit	Vector Labs	Cat# MB-9001
Flexi® Rabbit Reticulocyte Lysate System	Promega	Cat# L4540
Dual-Glo® Luciferase Assay System	Promega	Cat# E2920
Dual-Luciferase® Reporter Assay System	Promega	Cat# E1910
Deposited data		
Raw data files	This paper	Mendeley data: 10.17632/78kj389b6v.1
Crystal structure of the 2A-RNA pseudoknot complex	This paper	PDB: 9RVP
Small angle scattering data: 2A protein	This paper	SASBDB: SASDWQ5
Small angle scattering data: RNA	This paper	SASBDB: SASDWR5
Small angle scattering data: 2A-RNA complex	This paper	SASBDB: SASDWS5
Crystal structure of TMEV 2A protein	Hill <i>et al.</i> ⁵²	PDB: 7NBV
Crystal structure of EMCV 2A protein	Hill <i>et al.</i> ⁵³	PDB: 7BNY
Experimental models: Cell lines		
BHK-21	ATCC	CCL-10
HEK293T	ATCC	CRL-3216
Oligonucleotides		
See Table S1 for DNA oligonucleotides	This paper	N/A
See Table S2 for RNA oligonucleotides	This paper	N/A
Recombinant DNA		
pGEX6P2 vector containing wild type TMEV 2A (used for subcloning)	Napthine <i>et al.</i> ⁵¹	CH plasmid #73
pOPTnH vector containing inactive TMEV 2A _{R85A/R87A} (used for subcloning)	Hill <i>et al.</i> ⁵²	CH plasmid #33
pOPTHMBP TMEV wt 2A (used to make protein)	This paper	CH plasmid #219
pOPTHMBP TMEV 2A _{R85A/R87A} (used to make protein)	This paper	CH plasmid #423
pOPTHMBP TMEV 2A _{R7A} (used to make protein)	This paper	CH plasmid #323
pOPTHMBP TMEV 2A _{K24A} (used to make protein)	This paper	CH plasmid #324
pOPTHMBP TMEV 2A _{H26A} (used to make protein)	This paper	CH plasmid #325
pOPTHMBP TMEV 2A _{W89A} (used to make protein)	This paper	CH plasmid #326
pOPTnH EMCV wt 2A (used to make protein)	Hill <i>et al.</i> ⁵³	CH plasmid #31
pUCGW_TMEV-ns-SUMO (non-slip control for RRL <i>in vitro</i> assays)	This paper	CH plasmid #433
pUCGW_TMEV-ifc_SUMO (in-frame control for RRL <i>in vitro</i> assays)	This paper	CH plasmid #434
pUCGW_TMEV-wt_SUMO (wild type sequence for RRL <i>in vitro</i> assays)	This paper	CH plasmid #432
pUCGW_TMEV-S1-swap_SUMO (S1 swap test construct for RRL <i>in vitro</i> assays)	This paper	CH plasmid #428
pUCGW_TMEV-S1-AUstem_SUMO (S1 AU stem test construct for RRL <i>in vitro</i> assays)	This paper	CH plasmid #429
pUCGW_TMEV-S1-UAstem_SUMO (S1 UA stem test construct for RRL <i>in vitro</i> assays)	This paper	CH plasmid #430
pUCGW_TMEV-S1-deltaU19_SUMO (S1 ΔU19 test construct for RRL <i>in vitro</i> assays)	This paper	CH plasmid #431

pUCGW_TMEV-S2-plus1bp_SUMO (S2 +1 bp test construct for RRL <i>in vitro</i> assays)	This paper	CH plasmid #424
pUCGW_TMEV-S2-minus1bp_SUMO (S2 -1 bp test construct for RRL <i>in vitro</i> assays)	This paper	CH plasmid #425
pUCGW_TMEV-S2-minus2bp_SUMO (S2 -2 bp test construct for RRL <i>in vitro</i> assays)	This paper	CH plasmid #426
pUCGW_TMEV-S2-swap_SUMO (S2 swap test construct for RRL <i>in vitro</i> assays)	This paper	CH plasmid #427
pSGDlucV3.0	Gifted from Betty Chung ¹²¹	Loughran et al. ¹²⁰ addgene #119760
pSGDluc_TMEV-wt (wild type sequence for <i>vivo</i> dual luciferase assays)	This paper	CH plasmid #396
pSGDluc_TMEV-ifc (in-frame control for <i>in vivo</i> dual luciferase assays)	This paper	CH plasmid #397
pSGDluc_TMEV-S2-plus1bp (S2 + 1 bp for <i>in vivo</i> dual luciferase assays)	This paper	CH plasmid #398
pSGDluc_TMEV-S2-minus1bp (S2 - 1 bp for <i>in vivo</i> dual luciferase assays)	This paper	CH plasmid #399
pSGDluc_TMEV-S2-minus2bp (S2 - 2 bp for <i>in vivo</i> dual luciferase assays)	This paper	CH plasmid #400
pSGDluc_TMEV-S2-swap (S2 swap for <i>in vivo</i> dual luciferase assays)	This paper	CH plasmid #401
pSGDluc_TMEV-S1-swap (S1 swap for <i>in vivo</i> dual luciferase assays)	This paper	CH plasmid #402
pSGDluc_TMEV-S1-AUstem (S1 AU stem for <i>in vivo</i> dual luciferase assays)	This paper	CH plasmid #403
pSGDluc_TMEV-S1-UAstem (S1 UA stem for <i>in vivo</i> dual luciferase assays)	This paper	CH plasmid #404
pSGDluc_TMEV-S1-deltaU19 (S1 ΔU19 for <i>in vivo</i> dual luciferase assays)	This paper	CH plasmid #405
pCEP4 mammalian expression vector	Invitrogen	Cat# V04450
pCEP4-TMEV-2A-wt (for transient expression of wild type 2A <i>in vivo</i> dual luciferase assays)	This paper	CH plasmid #406
pCEP4-TMEV-2A _{R85A/R87A} (for transient expression of inactive 2A _{R85A/R87A} <i>in vivo</i> dual luciferase assays)	This paper	CH plasmid #407
Software and algorithms		
XIA2 automated pipeline	Winter ⁸⁹	https://xia2.github.io/index.html
XDS	Kabsch ⁹⁰	https://xds.mr.mpg.de/
AIMLESS	Evans and Murshudov ⁹¹	https://www.ccp4.ac.uk/
PHASER	McCoy <i>et al.</i> ⁹³	https://www.ccp4.ac.uk/
AlphaFold3	Abramson <i>et al.</i> ⁵⁷	https://alphafoldserver.com
COOT	Emsley <i>et al.</i> ⁹⁴	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
phenix.refine (v1.20.1)	Liebschner <i>et al.</i> ⁹⁵	https://www.phenix-online.org/
MolProbity	Chen <i>et al.</i> ⁹⁶	https://www.phenix-online.org/
PyMol	Schrödinger, LLC ⁹⁷	https://www.pymol.org/
ChimeraX	Meng <i>et al.</i> ⁹⁸	https://www.cgl.ucsf.edu/chimera/
3DNA webserver	Li <i>et al.</i> ⁹⁹	https://web.x3dna-dssr.org/

ATSAS software suite (v3.2.1)	Manalastas-Cantos <i>et al.</i> ¹⁰³	https://www.embl-hamburg.de/biosaxs/software.html
CHROMIXS	Panjikovich <i>et al.</i> ¹⁰¹	https://www.embl-hamburg.de/biosaxs/software.html
SASFLOW pipeline	Franke <i>et al.</i> ¹⁰²	https://www.embl-hamburg.de/biosaxs/software.html
Correlation Map (CorMap)	Franke <i>et al.</i> ¹⁰³	https://www.embl-hamburg.de/biosaxs/software.html
Primus	Manalastas-Cantos <i>et al.</i> ¹⁰⁴	https://www.embl-hamburg.de/biosaxs/software.html
<i>datvc</i>	Rambo and Tainer ¹⁰⁶	https://www.embl-hamburg.de/biosaxs/software.html
<i>datmow</i> module	Fischer <i>et al.</i> ¹⁰⁷	https://www.embl-hamburg.de/biosaxs/software.html
GNOM (v4.6)	Svergun ¹⁰⁹	https://www.embl-hamburg.de/biosaxs/software.html
MONSA (v1.44)	Svergun ¹¹⁰	https://www.embl-hamburg.de/biosaxs/software.html
MULCh	Whitten <i>et al.</i> ¹¹¹	https://smb-research.smb.usyd.edu.au/NCVWeb/
DAMMIN (v5.3)	Svergun ¹¹⁰	https://www.embl-hamburg.de/biosaxs/software.html
SREFLEX	Panjikovich & Svergun ¹¹³	https://www.embl-hamburg.de/biosaxs/software.html
CORAL (v1.1)	Petoukhov <i>et al.</i> ¹⁰⁵	https://www.embl-hamburg.de/biosaxs/software.html
CRY SOL (v2.8.3)	Svergun <i>et al.</i> ¹¹⁴	https://www.embl-hamburg.de/biosaxs/software.html
Anaconda Navigator	Anaconda	https://anaconda.org/
Jupyter Notebooks (v6.5.4)	Jupyter	https://jupyter.org/
FRETBursts (v0.8.3)	Ingargiola <i>et al.</i> ¹¹⁶	https://fretbursts.readthedocs.io/en/latest/#
FRET-restrained positioning and screening (FPS) tool (v1.1)	Kalinin <i>et al.</i> ¹¹⁹	https://www.mpc.hhu.de/software/fps
MO.Affinity Analysis software	NanoTemper Technologies	Cat# MO-S001A
ImageStudio 6.1	LICORbio	https://www.licorbio.com/image-studio
Spyder (v5.4.3)	Spyder IDE	https://docs.spyder-ide.org/current/index.html
SciPy (v1.10.1)	SciPy	https://docs.scipy.org/doc/scipy-1.10.1/index.html
InkScape (v1.4.2)	Inkscape	https://inkscape.org/

Other		
Ni Sepharose™ 6 Fast Flow histidine-tagged protein purification resin	Cytiva	Cat# 17531802
HiTrap™ Heparin HP affinity columns	Cytiva	Cat# 17040703
HiLoad™ 16/600 Superdex 75 pg	Cytiva	Cat# 28989333
SnakeSkin™ dialysis tubing 10K MWCO	Thermo Fisher	Cat# 88243
Amicon® Ultra Centrifugal Filter, 10 kDa MWCO	Merck	Cat# UFC901024
4–20% Mini-PROTEAN® TGX™ Precast Protein Gels	Bio Rad	Cat# 4561093
Protein LoBind® Tubes	Eppendorf	Cat# 0030108116
DNA LoBind® Tubes	Eppendorf	Cat# 0030108051
Monolith Premium Capillaries	NanoTemper	Cat# MO-K025
Cover slips, round #1, 30 mm	VWR	Cat# 631-0174
RNA Clean & Concentrator-5	Zymo Research	Cat# R1015

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Bacterial culture

All gene cloning, manipulation and plasmid propagation steps were performed in *Escherichia coli* Stellar cells (Takara) cultured in LB media with appropriate selection antibiotics.

Recombinant proteins: wild-type 2A, inactive 2A_{R85A/R87A}, 2A_{R7A}, 2A_{K24A}, 2A_{H26A} and 2A_{W89A} were expressed in *E.coli* BL21(DE3) pLysS cells grown in 2x TY media supplemented with appropriate antibiotics at 37°C, 200 rpm, until an optical density at 600 nm (OD₆₀₀) of 0.5 - 0.7 was reached. Expression was induced with 1.0 mM isopropyl β-d-1-thiogalactopyranoside (IPTG) and continued overnight (200 rpm, 21°C, 16 h). Bacteria were harvested (4000 × g, 4°C, 20 min), washed in cold 20 mM Tris-HCl pH 8.0 and stored at –20°C.

Cell lines

HEK293T cells (ATCC) were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS) at 37°C and 5% CO₂. BHK-21 cells (ATCC) were maintained in DMEM supplemented with 10% fetal bovine serum (FBS), 100 U/mL penicillin, 100 µg/mL streptomycin and 2 mM L-glutamine at 37°C and 5% CO₂.

METHOD DETAILS

Cloning

TMEV 2A cDNA was amplified by PCR from plasmid pOPT3G-based constructs (previously described in⁵²) and cloned into pOPTHMBP to introduce a TEV protease-cleavable N-terminal His₆ MBP tag. Mutant TMEV 2A proteins were prepared by mutagenesis PCR using the primers detailed in **Table S1**.

For *in vitro* translation assays, we designed reporter constructs with an EMCV internal ribosome entry site (IRES) to drive translation of a polypeptide including a 3x Myc tag, followed by nucleotides 4137-4307 of the TMEV genome (GenBank: NC_001366.1, corresponding to 35 codons upstream of the frameshift site, the slippery sequence and the stimulatory element), followed by the coding sequence for SUMO protein in the -1 frame. Constructs were designed

such that if elongation continues in the 0 frame the product is ~19 kDa, whereas shifting into the -1 frame enables translation to the end of SUMO, yielding a ~28 kDa product. To enable this, minor modifications were made to the wild-type TMEV sequence, as follows. The StopGo signal immediately before the frameshift site was inactivated by mutation of the NPGP motif to NPLV (insertion of CT between nucleotides 4229 and 4230, and deletion of C4234 and T4235). A -1 frame stop codon (UAA) within the stimulatory element was removed by introducing the T4274A mutation. All further S1 and S2 mutations were made in this context. The non-slippery (ns) control mRNA contains a mutated slippery sequence to prevent shifting into the -1 frame, and the in-frame control (ifc) mRNA introduces another nucleotide after the stimulatory element, to simulate the effect of 100% -1 frameshifting (i.e. by bringing the SUMO coding sequence into the 0 frame). These sequences were synthesised *de novo* and subcloned into the pUC-GW-Amp vector (GENEWIZ, Azenta Life Sciences).

To test frameshifting in cells, DNA sequences (nucleotides 4137-4307 of the TMEV genome, described above) corresponding to the modified wild-type sequence, mutants and controls, were amplified by PCR using primers in **Table S1** and subcloned into the pSGDlucV3.0¹²⁰ dual luciferase reporter plasmid (a gift from Betty Chung¹²¹) by restriction digest and ligation at the *PspXI*/*Bgl*III sites, to introduce sequences of interest between an upstream *Renilla* luciferase (R-luc) in the 0 frame, and a downstream Firefly luciferase (F-luc) in the -1 frame. F-Luc is only produced by a -1 frameshift, and the in-frame control (ifc) construct introduces another nucleotide after the stimulatory element, to simulate the effect of 100% -1 frameshifting (i.e. by bringing the F-luc coding sequence into the 0 frame).

Wild-type and inactive 2A_{R85A/R87A} were cloned into the pCEP4 (Invitrogen) mammalian expression vector by restriction digest and ligation at the *Xho*I/*Bam*HI sites. All constructs were verified by sanger sequencing.

Protein purification

TMEV 2A protein purification was performed as previously described⁵². Cell pellets from 2.0 L of *E.coli* BL21(DE3) pLysS cells were thawed in 300 mL of lysis buffer (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 10 mM imidazole supplemented with 50 µg/mL DNase I and EDTA-free proteases) and lysed using a cell disruptor (Constant Systems, 27 kPSI, 4°C). The insoluble fraction was removed by centrifugation (39,000 × g, 40 min, 4.0°C) and discarded. Supernatant was incubated (1 h, 4°C) with 4.0 mL Ni Sepharose 6 Fast Flow resin (Cytiva) pre-equilibrated in the same buffer. The resin was washed three times with 200 mL buffer (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 10 mM imidazole) in a gravity column. The protein was batch eluted with 20 mL of 50 mM Tris-HCl pH 8.0, 300 mM NaCl, 300 mM imidazole. His₆-tagged TEV protease (50 µg/mL, prepared in-house) was added to the eluate and was dialysed (3.5K MWCO, 4°C, 16 h) against 2.0 L buffer (20 mM Tris-HCl pH 7.4, 400 mM NaCl, 1.0 mM DTT). Dialysed proteins were subject to heparin-affinity chromatography to remove nucleic acids and the His₆-MBP solubility tag. Samples were loaded onto a 10 mL HiTrap Heparin column (Cytiva) at 1 mL/min, washed twice with buffer A (20 mM Tris-HCl pH 7.4, 400 mM NaCl, 1.0 mM DTT) and eluted with a 0% → 50% gradient of buffer B (20 mM Tris-HCl pH 7.4, 1.0 M NaCl, 1.0 mM DTT) over 8 CV before stepping to 100% buffer B over 10 CV. Fractions containing the 2A protein were pooled and concentrated in an Amicon Ultra centrifugal filter unit (10K MWCO, 4000 × g) prior to size exclusion chromatography (Superdex 75 16/600, 20 mM Tris-HCl pH 7.4, 1.0 M NaCl, 1.0 mM DTT). Purity was assessed by 4-20% gradient SDS-PAGE. Purified protein was either used immediately or was concentrated (~3.5 mg/mL, 220 µM), snap-frozen in liquid nitrogen and stored at -80°C.

EMCV 2A protein purification was performed as previously described⁵³ with minor modifications. Cell pellets from 2.0 L of *E.coli* BL21(DE3) pLysS cells were thawed in 300 mL of lysis buffer (50 mM Tris-HCl pH 8.0, 500 mM NaCl, 30 mM imidazole pH 8.0 supplemented with 50 µg/mL DNase I and EDTA-free proteases) and lysed using a cell disruptor (Constant Systems, 27 kPSI, 4°C). The insoluble fraction was removed by centrifugation (39,000 × g, 40 min, 4.0°C) and discarded. Supernatant was incubated (1 h, 4°C) with 4.0 mL Ni Sepharose 6 Fast Flow resin (Cytiva) pre-equilibrated in the same buffer. The resin was washed five times with 200 mL buffer (50 mM Tris-HCl pH 8.0, 500 mM NaCl, 30 mM imidazole pH 8.0) in a gravity column. The protein was batch eluted with 20 mL of 50 mM Tris-HCl pH 8.0, 500 mM NaCl, 300 mM imidazole and dialysed (3.5K MWCO, 4°C, 16 h) against 2.0 L buffer (50 mM Tris-HCl pH 8.0, 1.0 M NaCl, 1.0 mM DTT). EMCV 2A protein was concentrated in an Amicon Ultra centrifugal filter unit (10K MWCO, 4000 × g) prior to size exclusion chromatography (Superdex 75 16/600, 10 mM Tris-HCl pH 7.4, 1.0 M NaCl, 1.0 mM DTT). Purity was assessed by 4-20% gradient SDS-PAGE. Purified protein was concentrated (~4 mg/mL, 225 µM), snap-frozen in liquid nitrogen and stored at -80°C.

2A-RNA complex formation and crystallisation

RNA oligonucleotides (oligo 1, detailed in **Table S2**) were reconstituted in 10 mM Tris-HCl pH 7.4, 500 mM NaCl to 2.2 mg/mL (200 µM) determined by measuring absorbance at 260 nm using a Nanodrop spectrophotometer (Thermo). 2A-RNA complex was formed by diluting 2A protein to a final concentration of 500 mM NaCl and adding a 1:1 molar ratio of RNA. Complexes were incubated at room temperature for 20 min prior to buffer exchanging into 10 mM HEPES pH 7.4, 100 mM NaCl, 1.0 mM MgCl₂, 1.0 mM DTT and concentrating using an Amicon Ultra centrifugal filter unit (10K MWCO, 4000 × g, 10°C). Complexes were then subjected to size exclusion chromatography using Superdex 75 16/600 (Cytiva) at 10°C in 10 mM HEPES pH 7.4, 100 mM NaCl, 1.0 mM MgCl₂, 1.0 mM DTT. Complex formation was assessed by SDS-PAGE and urea-PAGE. Fractions containing the 2A-RNA complex were pooled and concentrated (~7.5 mg/mL, 693 µM) as assessed by absorbance at 260 nm. Sitting-drop vapor diffusion experiments were set-up in 96-well MRC plates with 50 µL reservoir solution, 150 nL complex and 150 nL crystallisation buffer at 21°C. Diffraction quality crystals grew in 0.2 M ammonium acetate, 0.1 M Bis-Tris pH 5.9, 19% (w/v) PEG 3350. Ten days after plating, crystals were cryo-protected by the addition of 0.2 µL crystallisation buffer supplemented with 20% (v/v) glycerol, prior to harvesting in nylon loops and flash-cooling by plunging into liquid nitrogen.

X-ray data collection, structure determination, refinement and analysis

Datasets of 3600 images were recorded at Diamond Light Source, beamline I04 ($\lambda = 0.9537$ Å) on an Eiger2 XE 16M detector (Dectris), using 100% transmission, an oscillation range of 0.10° and an exposure time of 0.02s per image. Data were processed with XIA2 automated pipeline⁸⁹ using XDS⁹⁰ for indexing and integration and AIMLESS⁹¹ for scaling and merging. Processing and refinement statistics are detailed in **Table 1**. Resolution cut-off was decided by CC_{1/2} values ≥ 0.5 and an $I/\sigma(I) \geq 2$ in the highest resolution shell⁹². Phasing was done by molecular replacement in PHASER⁹³ using two sequential search models. First a partial solution was obtained by placing the coordinates of the apo TMEV 2A protein (PDB: [7NBV](#)), followed by searching the unit cell for an idealised 7 bp A-form RNA helix (corresponding to S2) generated by AlphaFold3⁵⁷. The model was completed manually by iterative cycles of model building using COOT⁹⁴ and refinement with phenix.refine⁹⁵. MolProbity⁹⁶ was used to

assess model geometry through the refinement process. Structural figures (cartoon, stick and surface representations) were rendered in PyMOL (Schrödinger, LLC)⁹⁷ and ChimeraX⁹⁸. Analysis of the RNA pseudoknot geometry was performed using the 3DNA webserver⁹⁹.

SEC-SAXS

All small angle X-ray scattering (SAXS) experiments were performed coupled to size-exclusion chromatography (SEC-SAXS) at the EMBL P12 beamline at PETRA III (DESY; Hamburg, Germany)¹⁰⁰. Scattering data were collected on a Pilatus 6M detector at a sample-detector distance of 3 m and at a wavelength of $\lambda = 0.123983$ nm ($I(s)$ vs s , where $s = 4\pi\sin\theta/\lambda$, and 2θ is the scattering angle). For 2A protein-only samples, it was necessary to use high salt to prevent aggregation. 50–100 μ L protein at 4.0 mg/ml in 20 mM HEPES pH 7.4, 1.0 M KCl, 2.0% v/v glycerol, 1.0 mM DTT was injected at a 0.35 ml/min flow rate onto a GE Superdex 75 Increase 5/150 column at 20°C. 143 successive 0.25 second frames were collected through two consecutive SEC-SAXS runs, spanning the sample elution peaks of the protein. For RNA-only samples, 50 μ l at 7.1 mg/mL in 20 mM HEPES pH 7.4, 400 mM KCl, 2.0% v/v glycerol, 1.0 mM DTT was injected as above; 57 successive 0.25 second frames were collected through the SEC-elution peak. The 2A-RNA complex was prepared at 8.0 mg/mL in an identical buffer and injected as above; data were collected over 41 successive 0.25 second frames. Further details are provided in **Table 2**.

Initial data processing was carried out using CHROMIXS¹⁰¹ and azimuthal averaging from 2D to 1D was performed via the SASFLOW pipeline¹⁰². Buffer frames were evaluated for statistical similarity using all pairwise Correlation Map (CorMap)¹⁰³ p-values, applying a significance threshold (α) of 0.01. Statistically equivalent frames were then averaged to produce the final buffer scattering profile, which was subtracted from the corresponding macromolecular elution peaks. Subtracted data blocks that showed a consistent radius of gyration (R_g) values across the elution profile — assessed via the Guinier approximation — were scaled, compared using CorMap, and averaged to yield the final reduced 1D scattering profiles. All CorMap analyses were conducted using PrimusQT within the ATSAS software suite¹⁰⁴. Molecular weight estimations were derived using multiple approaches available in ATSAS, including the Porod volume via *datporod*¹⁰⁵, volume-of-correlation analysis via *datvc*¹⁰⁶, the *datmow* module¹⁰⁷ and the Bayesian consensus method¹⁰⁸. Indirect inverse Fourier transform of the SAXS data was performed using GNOM¹⁰⁹ to compute the real-space pair-distance distribution function, $p(r)$ versus r , from which the radius of gyration (R_g) and maximum particle dimension (D_{max}) were obtained. Further details are provided in **Table 2**.

Ab initio dummy atom bead modelling/shape reconstruction was performed using MONSA¹¹⁰, that takes into account differences in the average X-ray scattering length density of protein compared to RNA. Briefly, the 2A protein and RNA-only samples were modelled to their individual scattering profiles as single-phase particles, where the average excess X-ray scattering contrast of each phase was set to 2.413×10^{10} cm⁻² (2A protein) or 5.580×10^{10} cm⁻² (RNA) calculated relative to the average scattering length density of the respective sample buffers using MULCh¹¹¹. Target phase volumes were set to 21.4 nm³ (2A protein) and 11.2 nm³ (RNA), respectively, based on the amino acid or RNA composition¹¹¹. The 2A-RNA complex was modelled as a two-phase particle against the 2A protein and 2A-RNA complex SAXS datasets in parallel, using the scattering contrasts and target volumes described above for the protein and RNA phases. MONSA was run five times for each particle reconstruction to qualitatively evaluate shape/volume consistency. Additional single-phase dummy atom modelling for the 2A protein or the 2A-RNA complex was performed using DAMMIN¹¹⁰ using the output from GNOM as input. Following standard protocols¹¹² DAMMIN was run 10 times,

the individual models spatially aligned and averaged to develop a consensus shape reconstruction for either sample.

Atomistic models for the RNA were developed using a combination of AlphaFold 3 (AF3)⁵⁷ and normal-mode analysis. The RNA sequence was used as input into AF3 to generate five predicted RNA structures. These individual models underwent subsequent SREFLEX¹¹³ normal mode analysis, where the RNA structures were divided up into three rigid bodies that underwent repositioning relative to each other minimizing the fit to the RNA SAXS data (nucleotide ranges: 5'-CCACAA-3'; 5'-GGUGC-3'; 5'-GGUGCUAACUAAAUCCCUAGCACCCC-3'). Of the five initial AF3 models it was found that AF3 model 4 (AF3-4) acted as the best template with respect to generating a fit to the SAXS data after SREFLEX analysis. The post SREFLEX AF3-4 model underwent a second round of SREFLEX modelling to further evaluate the position/conformation of the first six nucleotides that appear as a 5'-overhang and not hybridized within the core RNA stem-loop. The SREFLEX AF3-4 model was divided into two rigid-bodies (5'-CCACAA-3'; 5'-GGUGCGGUGCUAACUAAAUCCCUAGCACCCC-3') where the first 5'-CCACAA-3' region underwent systematic manual rotations/translations to develop seven new SREFLEX AF3-4 based input templates, or alternatively these six 5'-nucleotides were swapped-out with the alternate AF3-predicted structures of the 5'-CCACAA-3' region derived from AF models 0, 1, 2 and 3. SREFLEX was performed on the 11 new input RNA templates to develop representative individual RNA structures that fit the SAXS data.

CORAL rigid body modelling^{104,105} was used to generate models of the 2A protein and 2A-RNA complex. In both cases, the structural templates of the protein and RNA components were obtained from the X-ray crystal structure of the 2A-RNA complex (this study, PDB 9RVP). Unresolved RNA nucleotides (U31, A32 and A33) of the 2A-RNA complex were built using COOT⁹⁴, by positioning a stereo-chemically ideal UAA trinucleotide between C30 and A34, linking the phosphodiester backbone manually and performing real-space regularisation of the three modelled nucleotides using the 2Fo-Fc crystallographic map. The unresolved/likely flexible N- and C-terminal extensions of the 2A protein component were represented as dummy amino acids (N-terminus: GPLGSPAS; C-terminus: HDVEMNPGEFPGRLEPHRD). CORAL was run five times against the respective 2A protein or 2A-RNA complex SAXS data to obtain a cohort of protein-alone or protein/RNA structures with alternative spatial positions of the 2A protein N- and C-terminal extensions. The final fit to the SAXS data for all atomistic models was performed using CRY SOL v2.8.3¹¹⁵ (30 harmonics, 256 points, constant enabled, average solvent electron density = 0.3416 e/Å³) and assessed by the reduced χ^2 test or CorMap p-value. Additional details are provided in **Table 2**. All subtracted and un-subtracted SEC-SAXS 1D data frames, R_g correlations through the respective SEC-SAXS peaks, as well as MONSA, DAMMIN, SREFLEX and CORAL models and the corresponding model fits are made available in the Small Angle Scattering Biological Databank (SASBDB)¹¹⁵ under the accession codes: SASDWQ5 (2A protein), SASDWR5 (RNA) and SASDWS5 (2A-RNA complex).

smFRET data acquisition and analysis

TMEV 'end reporter' (5' Cy5, 3' Cy3) and 'internal reporter' (5' Cy5, U31 Cy3; oligos 2 and 3 detailed in **Table S2**) RNA oligonucleotides (IDT) were resuspended to 100 μ M in nuclease-free water and stored at -80 °C prior to use. TMEV mutant 'end reporter' RNAs (oligos 16-19, detailed in **Table S2**) and EMCV 'end reporter' RNAs (oligos 20 and 21 detailed in **Table S2**) were prepared as described above. Samples were diluted to ~12.5 pM RNA in 10 mM Tris-HCl pH 7.4, 150 mM NaCl, 1.0 mM DTT, 1.0 mM MgCl₂ with 0.1 mg/mL BSA. For

measurements with 2A protein, 0.5 μM of TMEV or EMCV 2A protein was added to the buffer above prior to RNA dilution as appropriate. smFRET data on freely diffusing molecules was acquired using a confocal-based single-molecule FRET spectrometer (EI-FLEX, Exciting Instruments) equipped with 520 and 638 nm laser lines operating under alternating laser excitation (ALEX). The ALEX cycle period was 100 μs (45 μs donor excitation, 5 μs lasers off, 45 μs acceptor excitation, 5 μs lasers off). 50 μL of sample was placed in a sealed coverslip (#1) chamber and data recorded for 2 hr at ambient temperature.

Triplicate measurements from independent experiments were collected and data combined. Analysis was performed in Anaconda 3 with Jupyter Notebooks using FRETbursts¹¹⁶ (v 0.8.3) as described previously¹¹⁷. Briefly, apparent smFRET efficiencies, defined as $I_A/(I_A+I_D)$, where I_D and I_A are the donor and sensitised acceptor emission intensities under 520 nm excitation were corrected with all four correction parameters as described previously¹¹⁸. For our system correction parameters where: $\alpha = 0.08319$, $\delta = 0.06962$, $\gamma = 1.12903$, $\beta = 1.07552$. Molecular stoichiometries, S , were evaluated via $(I_D + I_A) / (I_D + I_A + I_A^*)$, where I_A^* is the acceptor emission intensity upon direct excitation. After background estimation and subtraction, bursts were identified using an all photon sliding window algorithm as previously described¹¹⁶ with a photon rate of 10 photons and a minimum photon rate six times greater than the background rate. A burst size threshold of 10 and 20 photons was applied for the donor and acceptor channels respectively. Multimer events were removed from the donor channel by applying an upper limit threshold of 100 photons. Generally, 54% of molecules were donor only, 2% acceptor only and 44% doubly labelled. A dual channel burst search was used to extract doubly labelled bursts from each acquisition and find E and S with one or two Gaussian fits. The apparent donor-acceptor distance was calculated by $R_{(E)} = R_0(E^{-1} - 1)^{\frac{1}{6}}$ where R_0 is 53 Å. We assumed random dipole orientation $\kappa^2 = \frac{2}{3}$ when approximating distances based on the peak FRET efficiency positions. Burst variance analysis (BVA) was performed as previously described⁶⁰ and implemented in Jupyter Notebooks using FRETbursts¹¹⁶. Bursts were divided into sub-bursts ($N = 7$ photons) and the standard deviation calculated for each sub-burst population over the time period of individual bursts. Accessible volume (AV) modelling was performed using the FRET-restrained positioning and screening (FPS) tool¹¹⁹. The three radii model was used with dye parameters: $R_1 = 11$ Å, $R_2 = 3$ Å and $R_3 = 1.5$ Å for Cy5 and $R_1 = 6.8$ Å, $R_2 = 3$ Å and $R_3 = 1.5$ Å for Cy3. Both dyes were modelled with a linker length of 14 Å and linker width of 4.5 Å. Approximate distances from dye mean points were measured in PyMOL⁹⁷.

Electrophoretic Mobility Shift Assay (EMSA)

RNA oligonucleotides (oligos 4-15, detailed in **Table S2**) were reconstituted in nuclease-free water to 100 μM . RNAs were labelled at the 5' end with Cy5-maleimide (GE Healthcare) or AzDye647-maleimide (Vector Labs) using the 5' EndTag kit (Vector Labs) according to the manufacturer's instructions, prior to phenol:chloroform extraction, ethanol precipitation and aqueous resuspension. The concentration of RNA was determined by measuring absorbance at 260 nm using a Nanodrop spectrophotometer (Thermo). 10 μL binding reactions contained 1.0 μL 500 nM fluorescently labelled RNA, 1.0 μL two-fold serially diluted 2A protein at concentrations of 0-12.5 μM in 10 mM Tris-HCl pH 7.4, 1.0 M NaCl, 1.0 mM DTT, 5.0 μL 2 $\times$ buffer (20 mM Tris-HCl pH 7.4, 80 mM NaCl, 4.0 mM magnesium acetate, 2.0 mM DTT, 10% v/v glycerol and 0.02% w/v bromophenol blue) and 3.0 μL distilled water. Final concentrations in the binding reactions were 50 nM RNA, 1 $\times$ buffer, a salt concentration of ~140 mM and 2A protein concentrations of 0-1.25 μM . Samples were incubated at 37°C for 20 mins prior to

analysis by non-denaturing 10% acrylamide/TBE PAGE (30 mins, 200 V constant). Gels were imaged with a Typhoon-5 using the 532 and 635 nm lasers/570BP20 and 670BP30 filters.

Microscale Thermophoresis (MST)

Two-fold dilution series of TMEV 2A wild-type protein or mutants were prepared in SEC buffer (20 mM Tris-HCl pH 7.4, 1.0 M NaCl, 1.0 mM DTT) producing 2A concentrations ranging from 50 μ M to 1.55 nM. For each binding experiment, fluorescently labelled RNA (oligos 4-15, detailed in **Table S2**) were diluted to 83.3 nM in MST buffer (20 mM Tris-HCl pH 7.4, 1.67 mM MgCl₂, 1.0 mM DTT, 0.083% Tween-20) supplemented with 0.167 mg/ml bovine liver tRNA (Sigma). For the measurement, 2A ligand dilutions were mixed with labelled RNA in a 3:2 ratio, for a final concentration of 50 nM labelled RNA in 20 mM Tris-HCl pH 7.4, 400 mM NaCl, 1.0 mM MgCl₂, 1.0 mM DTT, 0.05% Tween-20, 0.1 mg/ml tRNA and a 2A concentration series from 20 μ M to 0.61 nM. The reaction was mixed by pipetting, incubated for at least 5 min at room temperature and then held at 4 °C. Capillary forces were used to load the samples into Monolith NT.115 Premium Capillaries (NanoTemper Technologies). Measurements were performed using a Monolith NT.115 instrument (NanoTemper Technologies) at an ambient temperature of 22 °C. Measurements were taken at 20-100% LED power and high MST power. An initial fluorescence scan was performed across the capillaries to determine the sample quality and afterwards 16 subsequent thermophoresis measurements were performed. Data analysis was carried out with an MST on-time of 1.5 seconds to determine binding affinities (MO.Affinity Analysis software, NanoTemper Technologies). Data were fitted to the K_D model (single-site binding) using MO.Affinity Analysis software (NanoTemper) using the following equation:

$$f(C_{ligand}) = Unbound + (Bound - Unbound) \times \frac{C_{ligand} + C_{target} + K_D - \sqrt{(C_{ligand} + C_{target} + K_D)^2 - 4 \cdot C_{ligand} \cdot C_{target}}}{2C_{target}}$$

Where $f(C_{ligand})$ is the normalised fluorescence value at a given ligand concentration C_{ligand} . *Unbound* is the normalised fluorescence signal of the target alone, *Bound* is the normalised fluorescence signal of the complex. K_D is the dissociation constant and C_{target} is the final concentration of the target in the assay.

In vitro translation assays

Plasmids were linearised with NdeI (NEB) prior to run off transcription with T7 RNA polymerase (prepared in-house). Messenger RNAs were recovered by phenol:chloroform extraction (1:1 v/v) and concentrated by ethanol precipitation before aqueous resuspension. Further purification was performed using the RNA clean and concentrator-5 kit (Zymo Research) as per manufacturer's instructions. mRNA purity was assessed by gel electrophoresis and concentration determined by absorbance at 260 nm.

Messenger RNAs were translated in Flexi® rabbit reticulocyte lysate (RRL) (Promega). Typical reactions (10 μ L) were composed of 66% (v/v) RRL, 25 μ M amino acids (lacking methionine) and 0.2 MBq [³⁵S]-methionine (Hartmann Analytic). The mixture was pre-equilibrated for 10 mins at 30 °C to allow for tRNA charging, prior to addition of the mRNAs and 2A protein to start the reaction. Translation reactions were conducted for 1 hr (30 °C), then stopped by addition of 100 μ g/ μ L RNase A in 10 mM EDTA pH 8.0 (20 min, room temperature), followed by addition of 10 volumes of 2x Laemmli's sample buffer, and heating (5 min, 95°C). Translation products were assessed by 15% SDS-PAGE. Dried gels were exposed to a

Storage Phosphor Screen (PerkinElmer) and the screen scanned in a Typhoon-5 (Cytiva) using Phosphor imaging mode. Frameshift assays were performed in triplicate unless otherwise stated.

Dual luciferase assays in cells

Adherent BHK-21 and HEK293T cells were transfected after seeding in 24-well plates to achieve 50% confluency. BHK-21 cells were transfected with 625 ng total plasmids at equimolar concentrations (250 ng pSGDlucV3.0 reporter and 325 ng pCEP4 expressing 2A) using TransIT-LT1 transfection reagent (Mirus Bio). HEK293T cells were transfected with 1 µg total plasmids in 1:1 ratio of pSGDlucV3.0 reporter to pCEP4 2A using PEI transfection reagent (MedChemExpress) as per manufactures instructions. BHK-21 cells were harvested 36 hpt, whereas HEK293T cells were harvested 24 hpt. Cells were washed with PBS and lysed with 1× passive lysis buffer (Promega), prior to luciferase activity measurement using either the Dual-glo® (HEK293T) or the Dual-luciferase® (BHK-21) reporter assay kits (Promega) according to the manufacturer's instructions. Measurements were recorded using a GloMax Navigator microplate luminometer (Promega). The entire assay, from transfection to measurement, was performed in quadruplicate (BHK-21 cells) or triplicate (HEK293T cells).

QUANTIFICATION AND STATISTICAL ANALYSIS

Crystallographic calculations (integration, scaling, merging) were performed as described in the methods text using the default software parameters unless otherwise stated. Processing and refinement statistics are detailed in **Table 1**.

SEC-SAXS calculations are described in methods text and detailed in **Table 2**. Additional information can be found in the Small Angle Scattering Biological Data Bank (SASBDB) under accession codes SASDWQ5 (2A protein only), SASDWR5 (RNA only) and SASDWS5 (2A-RNA complex).

smFRET data analysis was performed as described in methods text. Example Jupyter Notebooks for data analysis are deposited in Mendeley Data **10.17632/78kj389b6v.1**. FRET efficiencies are presented in the text as the mean ± SD. Data was collected and combined from three independent replicates. n indicates the number of bursts analysed after a dual channel burst search extracted doubly labelled molecules from each dataset.

For *in vitro* translation assays autoradiographs were quantified using ImageStudio 6.1 (LICORbio). The calculations of -1 frameshifting efficiency accounted for the difference in methionine content in the 0 and -1 frame products using the following equation:

$$-1 \text{ frameshift efficiency (\%)} = \frac{(IFS \div MetFS)}{((IS \div MetS) + (IFS \div MetFS))} \times 100$$

Where the number of methionines in the 0 frame and -1 frame products are denoted as MetS and MetFS respectively. The densitometry values for the same products are denoted by IS and IFS respectively. Frameshift assays performed in triplicate unless otherwise stated, where data is presented as the mean ± SD.

Frameshifting efficiencies in cells were determined by calculating the relative *Renilla*/firefly luciferase activities from each test construct and dividing by the relative luciferase activity from in-frame control (ifc) constructs, as described in the equation below:

$$-1 \text{ frameshift efficiency (\%)} = \frac{(RLuc_{test} \div FLuc_{test})}{(RLuc_{ifc} \div FLuc_{ifc})} \times 100$$

A one-way ANOVA followed by Tukey HSD test was used to test for significant differences in the mean -1 frameshifting efficiency of each test construct compared to wild-type RNA with wild-type 2A.

Graphing, data and statistical analysis was performed in Python using Spyder (v5.4.3). For statistical analysis the python package SciPy (v1.10.1) was used. Further statistical details can be found in figure legends and in the tables. For all figure legends, n is indicative of the number of independent experimental replicates conducted.

References

1. Hill, C.H., and Brierley, I. (2023). Structural and functional insights into viral programmed ribosomal frameshifting. *Annu Rev Virol* 17, 18. <https://doi.org/10.1146/annurev-virology-111821-120646>.
2. Atkins, J.F., Loughran, G., Bhatt, P.R., Firth, A.E., and Baranov, P. V. (2016). Ribosomal frameshifting and transcriptional slippage: From genetic steganography and cryptography to adventitious use. *Nucleic Acids Res.* 44, 7007–7078. <https://doi.org/10.1093/nar/gkw530>.
3. Korniy, N., Samatova, E., Anokhina, M.M., Peske, F., and Rodnina, M. V. (2019). Mechanisms and biomedical implications of –1 programmed ribosome frameshifting on viral and bacterial mRNAs. *FEBS Lett.* 593, 1468–1482. <https://doi.org/10.1002/1873-3468.13478>.
4. World Health Organization (2025). World health statistics 2025: monitoring health for the SDGs, sustainable development goals.
5. World Health Organization (2023). Coronavirus disease (COVID-19) dashboard. <https://data.who.int/>.
6. Quan, T.M., Thao, T.T.N., Duy, N.M., Nhat, T.M., and Clapham, H.E. (2020). Estimates of the global burden of Japanese encephalitis and the impact of vaccination from 2000-2015. *Elife* 9, 1–187. <https://doi.org/10.7554/eLife.51027>.
7. Neumann, E.J., Kliebenstein, J.B., Johnson, C.D., Mabry, J.W., Bush, E.J., Seitzinger, A.H., Green, A.L., and Zimmerman, J.J. (2005). Assessment of the economic impact of porcine reproductive and respiratory syndrome on swine production in the United States. *J Am Vet Med Assoc* 227, 385–392. <https://doi.org/10.2460/javma.2005.227.385>.
8. Jacks, T., and Varmus, H.E. (1985). Expression of the rous sarcoma virus pol gene by ribosomal frameshifting. *Science* (1979). 230, 1237–1242. <https://doi.org/10.1126/science.2416054>.
9. Jacks T, Power MD, Masiarz FR, Luciw PA, Barr PJ, and Varmus HE. (1988). Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. *Nature* 331, 280–283. <https://doi.org/10.1038/331280a0>.
10. Anokhina, V.S., and Miller, B.L. (2021). Targeting ribosomal frameshifting as an antiviral strategy: from HIV-1 to SARS-CoV-2. *Acc. Chem. Res.* 54, 3349–3361. <https://doi.org/10.1021/acs.accounts.1c00316>.
11. Matsumoto, S., Caliskan, N., Rodnina, M. V., Murata, A., and Nakatani, K. (2018). Small synthetic molecule-stabilized RNA pseudoknot as an activator for -1 ribosomal frameshifting. *Nucleic Acids Res.* 46, 8079–8089. <https://doi.org/10.1093/nar/gky689>.
12. Ritchie, D.B., Soong, J., Sikkema, W.K.A., and Woodside, M.T. (2014). Anti-frameshifting ligand reduces the conformational plasticity of the SARS virus pseudoknot. *J. Am. Chem. Soc.* 136, 2196–2199. <https://doi.org/10.1021/ja410344b>.
13. Takyar, S., Hickerson, R.P., and Noller, H.F. (2005). mRNA helicase activity of the ribosome. *Cell* 120, 49–58. <https://doi.org/10.1016/j.cell.2004.11.042>.

- 1024 14. Desai, V.P., Frank, F., Lee, A., Righini, M., Lancaster, L., Noller, H.F., Tinoco, I., and
1025 Bustamante, C. (2019). Co-temporal force and fluorescence measurements reveal a
1026 ribosomal gear shift mechanism of translation regulation by structured mRNAs. *Mol.*
1027 *Cell* 75, 1007-1019.e5. <https://doi.org/10.1016/j.molcel.2019.07.024>.
- 1028 15. Somogyi, P., Jenner, A.J., Brierley, I., and Inglist, S.C. (1993). Ribosomal pausing
1029 during translation of an RNA pseudoknot. *Mol Cell Biol* 13, 6931–6940.
1030 <https://doi.org/10.1128/mcb.13.11.6931-6940.1993>.
- 1031 16. Tu, C., Tzeng, T.H., and Bruenn, J.A. (1992). Ribosomal movement impeded at a
1032 pseudoknot required for frameshifting. *Proc Natl Acad Sci USA* 89, 8636–8640.
1033 <https://doi.org/10.1073/pnas.89.18.8636>.
- 1034 17. Lopinski, J.D., Dinman, J.D., and Bruenn, J.A. (2000). Kinetics of ribosomal pausing
1035 during programmed -1 translational frameshifting. *Mol Cell Biol* 20, 1095–1103.
1036 <https://doi.org/10.1128/MCB.20.4.1095-1103.2000>
- 1037 18. Kontos, H., Naphine, S., and Brierley, I. (2001). Ribosomal pausing at a frameshifter
1038 RNA pseudoknot is sensitive to reading phase but shows little correlation with
1039 frameshift efficiency. *Mol. Cell. Biol.* 21, 8657–8670.
1040 <https://doi.org/10.1128/mcb.21.24.8657-8670.2001>.
- 1041 19. Lin, Z., Gilbert, R.J.C., and Brierley, I. (2012). Spacer-length dependence of
1042 programmed -1 or -2 ribosomal frameshifting on a U6A heptamer supports a role for
1043 messenger RNA (mRNA) tension in frameshifting. *Nucleic Acids Res.* 40, 8674–8689.
1044 <https://doi.org/10.1093/nar/gks629>.
- 1045 20. Poulis, P., Peske, F., and Rodnina, M. V. (2023). The many faces of ribosome
1046 translocation along the mRNA: Reading frame maintenance, ribosome frameshifting
1047 and translational bypassing. *Biol. Chem.* 404, 755–767. [https://doi.org/10.1515/hsz-](https://doi.org/10.1515/hsz-2023-0142)
1048 [2023-0142](https://doi.org/10.1515/hsz-2023-0142).
- 1049 21. Chen, J., Petrov, A., Johansson, M., Tsai, A., O'leary, S.E., and Puglisi, J.D. (2014).
1050 Dynamic pathways of -1 translational frameshifting. *Nature* 512, 328–332.
1051 <https://doi.org/10.1038/nature13428>.
- 1052 22. Choi, J., O'Loughlin, S., Atkins, J.F., and Puglisi, J.D. (2020). The energy landscape of
1053 -1 ribosomal frameshifting. *Sci Adv* 6. <https://doi.org/10.1126/sciadv.aax6969>.
- 1054 23. Bock, L. V., Caliskan, N., Korniy, N., Peske, F., Rodnina, M. V., and Grubmüller, H.
1055 (2019). Thermodynamic control of -1 programmed ribosomal frameshifting. *Nat.*
1056 *Commun.* 10. <https://doi.org/10.1038/s41467-019-12648-x>.
- 1057 24. Liu, T., Kaplan, A., Alexander, L., Yan, S., Wen, J. Der, Lancaster, L., Wickersham,
1058 C.E., Fredrick, K., Noller, H., Tinoco, I., et al. (2014). Direct measurement of the
1059 mechanical work during translocation by the ribosome. *Elife* 3, 1–12.
1060 <https://doi.org/10.7554/eLife.03406>.
- 1061 25. Chen, X., Chamorro, M., Lee, S. I., Shen, L. X., Hinesl, J. V., Tinoco, I., and Varmus,
1062 H. E. (1995). Structural and functional studies of retroviral RNA pseudoknots involved
1063 in ribosomal frameshifting: nucleotides at the junction of the two stems are important
1064 for efficient ribosomal frameshifting. *EMBO J.* 14, 842–852.
- 1065 26. Su L, Chen L, Egli M, Berger JM, and Rich A (1999). Minor groove RNA triplex in the
1066 crystal structure of a ribosomal frameshifting viral pseudoknot. *Nat. Struct. Biol.* 6,
1067 285–295. <https://doi.org/10.1038/6722>.

- 1068 27. Nixon, P.L., Cornish, P. V., Suram, S. V., and Giedroc, D.P. (2002). Thermodynamic
1069 analysis of conserved loop-stem interactions in P1-P2 frameshifting RNA pseudoknots
1070 from plant Luteoviridae. *Biochemistry* **41**, 10665–10674.
1071 <https://doi.org/10.1021/bi025843c>.
- 1072 28. Pallan, P.S., Marshall, W.S., Harp, J., Jewett, F.C., Wawrzak, Z., Brown, B.A., Rich,
1073 A., and Egli, M. (2005). Crystal structure of a luteoviral RNA pseudoknot and model
1074 for a minimal ribosomal frameshifting motif. *Biochemistry* **44**, 11315–11322.
1075 <https://doi.org/10.1021/bi051061i>.
- 1076 29. Cornish, P. V., Hennig, M., and Giedroc, D.P. (2005). A loop 2 cytidine-stem 1 minor
1077 groove interaction as a positive determinant for pseudoknot-stimulated 1 ribosomal
1078 frameshifting. *Proc Natl Acad Sci USA* **102**, 12694–12699.
1079 <https://doi.org/10.1073/pnas.0506166102>
- 1080 30. Michiels, P.J.A., Versleijen, A.A.M., Verlaan, P.W., Pleij, C.W.A., Hilbers, C.W., and
1081 Heus, H.A. (2001). Solution structure of the pseudoknot of SRV-1 RNA, involved in
1082 ribosomal frameshifting. *J. Mol. Biol.* **310**, 1109–1123.
1083 <https://doi.org/10.1006/jmbi.2001.4823>.
- 1084 31. Cornish, P. V., Stammer, S.N., and Giedroc, D.P. (2006). The global structures of a
1085 wild-type and poorly functional plant luteoviral mRNA pseudoknot are essentially
1086 identical. *RNA* **12**, 1959–1969. <https://doi.org/10.1261/rna.199006>.
- 1087 32. Chen, X., Kang, H., Shen, L.X., Chamorro, M., Varmus, H.E., and Tinoco, I. (1996). A
1088 characteristic bent conformation of RNA pseudoknots promotes -1 frameshifting
1089 during translation of retroviral RNA. *J. Mol. Biol* **260**, 479–483.
1090 <https://doi.org/10.1006/jmbi.1996.0415>
- 1091 33. Shen, L.X., and Tinoco, I. (1995). The structure of an RNA pseudoknot that causes
1092 efficient frameshifting in mouse mammary tumor virus. *J. Mol. Biol* **247**, 963–978.
1093 <https://doi.org/10.1006/jmbi.1995.0193>
- 1094 34. Roman, C., Lewicka, A., Koirala, D., Li, N.S., and Piccirilli, J.A. (2021). The SARS-
1095 CoV-2 programmed -1 ribosomal frameshifting element crystal structure solved to
1096 2.09 Å using chaperone-assisted RNA crystallography. *ACS Chem. Biol.* **16**, 1469–
1097 1481. <https://doi.org/10.1021/acscchembio.1c00324>.
- 1098 35. Jones, C.P., and Ferré-D’amaré, A.R. (2022). Crystal structure of the severe acute
1099 respiratory syndrome coronavirus 2 (SARS-CoV-2) frameshifting pseudoknot. *RNA*
1100 **28**, 239. <https://doi.org/10.1261/rna.078825.121>.
- 1101 36. Zhang, K., Zheludev, I.N., Hagey, R.J., Haslecker, R., Hou, Y.J., Kretsch, R., Pintilie,
1102 G.D., Rangan, R., Kladwang, W., Li, S., et al. (2021). Cryo-EM and antisense
1103 targeting of the 28-kDa frameshift stimulation element from the SARS-CoV-2 RNA
1104 genome. *Nat. Struct. Mol. Biol.* **28**, 747–754. <https://doi.org/10.1038/s41594-021-00653-y>.
- 1106 37. Bhatt, P.R., Scaiola, A., Loughran, G., Leibundgut, M., Kratzel, A., Meurs, R., Dreos,
1107 R., O’Connor, K.M., McMillan, A., Bode, J.W., et al. (2021). Structural basis of
1108 ribosomal frameshifting during translation of the SARS-CoV-2 RNA genome. *Science*
1109 (1979). **372**, 1306–1313. <https://doi.org/10.1126/science.abf3546>.
- 1110 38. Ritchie, D.B., Foster, D.A.N., and Woodside, M.T. (2012). Programmed -1
1111 frameshifting efficiency correlates with RNA pseudoknot conformational plasticity, not

1112 resistance to mechanical unfolding. *Proc Natl Acad Sci USA* *109*, 16167–16172.
 1113 <https://doi.org/10.1073/pnas.1204114109>.

1114 39. Halma, M.T.J., Ritchie, D.B., and Woodside, M.T. (2021). Conformational shannon
 1115 entropy of mRNA structures from force spectroscopy measurements predicts the
 1116 efficiency of -1 programmed ribosomal frameshift stimulation. *Phys. Rev. Lett.* *126*.
 1117 <https://doi.org/10.1103/PhysRevLett.126.038102>.

1118 40. Halma, M.T.J., Ritchie, D.B., Cappellano, T.R., Neupane, K., and Woodside, M.T.
 1119 (2019). Complex dynamics under tension in a high-efficiency frameshift stimulatory
 1120 structure. *Proc. Natl. Acad. Sci. U. S. A.* *116*, 19500–19505.
 1121 <https://doi.org/10.1073/pnas.1905258116>.

1122 41. Yan, S., and Schlick, T. (2025). Heterogeneous and multiple conformational transition
 1123 pathways between pseudoknots of the SARS-CoV-2 frameshift element. *122*.
 1124 <https://doi.org/10.1073/pnas.2417479122>.

1125 42. Ludaic, M., and Elofsson, A. (2025). Limits of deep-learning-based RNA prediction
 1126 methods. Preprint, <https://doi.org/10.1101/2025.04.30.651414>
 1127 <https://doi.org/10.1101/2025.04.30.651414>.

1128 43. Wang, X., Xuan, Y., Han, Y., Ding, X., Ye, K., Yang, F., Gao, P., Goff, S.P., and Gao, G.
 1129 (2019). Regulation of HIV-1 Gag-Pol expression by shiftless, an inhibitor of
 1130 programmed -1 ribosomal frameshifting. *Cell* *176*, 625-635.e14.
 1131 <https://doi.org/10.1016/j.cell.2018.12.030>.

1132 44. Napthine, S., Hill, C.H., Nugent, H.C.M., and Brierley, I. (2021). Modulation of viral
 1133 programmed ribosomal frameshifting and stop codon readthrough by the host
 1134 restriction factor shiftless. *Viruses* *13*. <https://doi.org/10.3390/v13071230>.

1135 45. Zimmer, M.M., Kibe, A., Rand, U., Pekarek, L., Ye, L., Buck, S., Smyth, R.P., Cicin-
 1136 Sain, L., and Caliskan, N. (2021). The short isoform of the host antiviral protein ZAP
 1137 acts as an inhibitor of SARS-CoV-2 programmed ribosomal frameshifting. *Nat.*
 1138 *Commun.* *12*. <https://doi.org/10.1038/s41467-021-27431-0>.

1139 46. Li, Y., Treffers, E.E., Napthine, S., Tas, A., Zhu, L., Sun, Z., Bell, S., Mark, B.L., Van
 1140 Veelen, P.A., Van Hemert, M.J., et al. (2014). Transactivation of programmed
 1141 ribosomal frameshifting by a viral protein. *Proc. Natl. Acad. Sci. U. S. A.* *111*.
 1142 <https://doi.org/10.1073/pnas.1321930111>.

1143 47. Napthine, S., Treffers, E.E., Bell, S., Goodfellow, I., Fang, Y., Firth, A.E., Snijder, E.J.,
 1144 and Brierley, I. (2016). A novel role for poly(C) binding proteins in programmed
 1145 ribosomal frameshifting. *Nucleic Acids Res.* *44*, 5491–5503.
 1146 <https://doi.org/10.1093/nar/gkw480>.

1147 48. Loughran, G., Firth, A.E., and Atkins, J.F. (2011). Ribosomal frameshifting into an
 1148 overlapping gene in the 2B-encoding region of the cardiovirus genome. *Proc Natl*
 1149 *Acad Sci USA* *108*, 1111–1119. <https://doi.org/10.1073/pnas.1102932108>

1150 49. Napthine, S., Ling, R., Finch, L.K., Jones, J.D., Bell, S., Brierley, I., and Firth, A.E.
 1151 (2017). Protein-directed ribosomal frameshifting temporally regulates gene
 1152 expression. *Nat. Commun.* *8*. <https://doi.org/10.1038/ncomms15582>.

1153 50. Finch, L.K., Ling, R., Napthine, S., Olsper, A., Michiels, T., Lardinois, C., Bell, S.,
 1154 Loughran, G., Brierley, I., and Firth, A.E. (2015). Characterization of ribosomal

frameshifting in Theiler's murine encephalomyelitis virus. *J. Virol.* **89**, 8580–8589.
<https://doi.org/10.1128/jvi.01043-15>.

51. Napthine, S., Bell, S., Hil, C.H., Brierley, I., and Firth, A.E. (2019). Characterization of the stimulators of protein-directed ribosomal frameshifting in Theiler's murine encephalomyelitis virus. *Nucleic Acids Res.* **47**, 8207–8223.
<https://doi.org/10.1093/nar/gkz503>.

52. Hill, C.H., Cook, G.M., Napthine, S., Kibe, A., Brown, K., Caliskan, N., Firth, A.E., Graham, S.C., and Brierley, I. (2021). Investigating molecular mechanisms of 2A-stimulated ribosomal pausing and frameshifting in Theilovirus. *Nucleic Acids Res.* **49**, 11938–11958. <https://doi.org/10.1093/nar/gkab969>.

53. Hill, C.H., Pekarek, L., Napthine, S., Kibe, A., Firth, A.E., Graham, S.C., Caliskan, N., and Brierley, I. (2021). Structural and molecular basis for Cardiovirus 2A protein as a viral gene expression switch. *Nat. Commun.* **12**. <https://doi.org/10.1038/s41467-021-27400-7>.

54. Calnan B J, Tidor B, Biancalana S, Hudson D, and Frankel A D (1991). Arginine-mediated RNA recognition: the arginine fork. *Science* (1979). **142**, 34.
<https://doi.org/10.1126/science.252.5009.1167>.

55. Chavali, S.S., Cavender, C.E., Mathews, D.H., and Wedekind, J.E. (2020). Arginine forks are a widespread motif to recognize phosphate backbones and guanine nucleobases in the RNA major groove. *J. Am. Chem. Soc.* **142**, 19835–19839.
<https://doi.org/10.1021/jacs.0c09689>.

56. Jangra, R., Kukhta, T., Trant, J.F., and Sharma, P. (2024). Decoding the enigma of RNA-protein recognition: quantum chemical insights into arginine fork motifs. *Phys Chem Chem Phys* **26**, 28091–28100. <https://doi.org/10.1039/d4cp03987d>

57. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A.J., Bambrick, J., et al. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500.
<https://doi.org/10.1038/s41586-024-07487-w>.

58. Peselis, A., and Serganov, A. (2014). Structure and function of pseudoknots involved in gene expression control. *Wiley Interdiscip. Rev. RNA* **5**, 803–822.
<https://doi.org/10.1002/wrna.1247>.

59. Kapanidis, A.N., Lee, N.K., Laurence, T.A., Ren Dooze, S., Margeat, E., and Weiss, S. (2004). Fluorescence-aided molecule sorting: Analysis of structure and interactions by alternating-laser excitation of single molecules. *Proc Natl Acad Sci USA* **101**, 8936–8941. <https://doi.org/10.1073/pnas.0401690101>.

60. Torella, J.P., Holden, S.J., Santoso, Y., Hohlbein, J., and Kapanidis, A.N. (2011). Identifying molecular dynamics in single-molecule fret experiments with burst variance analysis. *Biophys. J.* **100**, 1568–1577. <https://doi.org/10.1016/j.bpj.2011.01.066>.

61. Napthine, S., Liphardt, J., Bloys, A., Routledge, S., and Brierley, I. (1999). The role of RNA pseudoknot stem 1 length in the promotion of efficient -1 ribosomal frameshifting. *J Mol Biol* **3**, 305–320. <https://doi.org/10.1006/jmbi.1999.2688>

62. Bidou L, Stahl G, Grima B, Liu H, Cassan M, and Rousset JP (1997). In vivo HIV-1 frameshifting efficiency is directly related to the stability of the stem-loop stimulatory signal. *RNA* **10**.

- 1199 63. Mouzakis, K.D., Lang, A.L., Vander Meulen, K.A., Easterday, P.D., and Butcher, S.E.
1200 (2013). HIV-1 frameshift efficiency is primarily determined by the stability of base pairs
1201 positioned at the mRNA entrance channel of the ribosome. *Nucleic Acids Res.* *41*,
1202 1901–1913. <https://doi.org/10.1093/nar/gks1254>.
- 1203 64. Moafinejad, S.N., De Aquino, B.R.H., Boniecki, M.J., Pandaranadar Jeyeram, I.P.N.,
1204 Nikolaev, G., Magnus, M., Farsani, M.A., Badepally, N.G., Wirecki, T.K., Stefaniak, F.,
1205 et al. (2024). SimRNAweb v2.0: a web server for RNA folding simulations and 3D
1206 structure modeling, with optional restraints and enhanced analysis of folding
1207 trajectories. *Nucleic Acids Res.* *52*, W368–W373. <https://doi.org/10.1093/nar/gkae356>.
- 1208 65. Bernard, C., Postic, G., Ghannay, S., and Tah, F. (2025). Has AlphaFold3 achieved
1209 success for RNA? *Acta Crystallogr. D Struct. Biol.* *81*, 49–62.
1210 <https://doi.org/10.1107/S2059798325000592>.
- 1211 66. Egli, M., Minasov, G., Li, S., and Rich, A. (2002). Metal ions and flexibility in a viral
1212 RNA pseudoknot at atomic resolution. *Proc Natl Acad Sci USA* *99*.
1213 <https://doi.org/10.1073/pnas.062055599>.
- 1214 67. Nixon, P.L., Rangan, A., Kim, Y.G., Rich, A., Hoffman, D.W., Hennig, M., and Giedroc,
1215 D.P. (2002). Solution structure of a luteoviral P1-P2 frameshifting mRNA pseudoknot.
1216 *J. Mol. Biol.* *322*, 621–633. [https://doi.org/10.1016/S0022-2836\(02\)00779-9](https://doi.org/10.1016/S0022-2836(02)00779-9).
- 1217 68. Pleij C W, Rietveld K, and Bosch L (1985). A new principle of RNA folding based on
1218 pseudoknotting. *Nucleic Acids Res.* *13*. <https://doi.org/10.1093/nar/13.5.1717>.
- 1219 69. Salvail, H., and Breaker, R.R. (2023). Riboswitches. *Current Biology* *33*, R343–R348.
1220 <https://doi.org/10.1016/j.cub.2023.03.069>.
- 1221 70. Breaker, R.R. (2018). Riboswitches and translation control. *Cold Spring Harb.*
1222 *Perspect. Biol.* *10*. <https://doi.org/10.1101/cshperspect.a032797>.
- 1223 71. Roth, A., Winkler, W.C., Regulski, E.E., Lee, B.W.K., Lim, J., Jona, I., Barrick, J.E.,
1224 Ritwik, A., Kim, J.N., Welz, R., et al. (2007). A riboswitch selective for the queuosine
1225 precursor preQ1 contains an unusually small aptamer domain. *Nat. Struct. Mol. Biol.*
1226 *14*, 308–317. <https://doi.org/10.1038/nsmb1224>.
- 1227 72. Nahvi, A., Sudarsan, N., Ebert, M.S., Zou, X., Brown, K.L., and Breaker, R.R. (2002).
1228 Genetic control by a metabolite binding mRNA. *Chem Biol* *9*, 1043–1049.
1229 [https://doi.org/10.1016/s1074-5521\(02\)00224-7](https://doi.org/10.1016/s1074-5521(02)00224-7).
- 1230 73. Grundy, F.J., and Henkin, T.M. (2002). The S box regulon: A new global transcription
1231 termination control system for methionine and cysteine biosynthesis genes in Gram-
1232 positive bacteria. *Mol. Microbiol.* *30*, 737–749. <https://doi.org/10.1046/j.1365-2958.1998.01105.x>.
- 1234 74. Winkler, W.C., Nahvi, A., Sudarsan, N., Barrick, J.E., and Breaker, R.R. (2003). An
1235 mRNA structure that controls gene expression by binding S-adenosylmethionine. *Nat.*
1236 *Struct. Biol.* *10*, 701–707. <https://doi.org/10.1038/nsb967>.
- 1237 75. Sudarsan, N., Wickiser, J.K., Nakamura, S., Ebert, M.S., and Breaker, R.R. (2003). An
1238 mRNA structure in bacteria that controls gene expression by binding lysine. *Genes*
1239 *Dev.* *17*, 2688–2697. <https://doi.org/10.1101/gad.1140003>.
- 1240 76. Mandal, M., Lee, M., Barrick, J.E., Weinburg, Z., Emilsson, G.M., Ruzzo, W.L., and
1241 Breaker, R.R. (2004). A glycine-dependent riboswitch that uses cooperative binding to

control gene expression. *Science* (1979). *306*, 275–279.
<https://doi.org/10.1126/science.1100829>.

77. Klähn, S., Bolay, P., Wright, P.R., Atilho, R.M., Brewer, K.I., Hagemann, M., Breaker, R.R., and Hess, W.R. (2018). A glutamine riboswitch is a key element for the regulation of glutamine synthetase in cyanobacteria. *Nucleic Acids Res.* *46*, 10082–10094. <https://doi.org/10.1093/nar/gky709>.

78. Grundy, F.J., and Henkin, T.M. (1993). tRNA as a positive regulator of transcription antitermination in *B. subtilis*. *Cell* *74*, 475–482. [https://doi.org/10.1016/0092-8674\(93\)80049-k](https://doi.org/10.1016/0092-8674(93)80049-k).

79. de Jesus, V., Qureshi, N.S., Warhaut, S., Bains, J.K., Dietz, M.S., Heilemann, M., Schwalbe, H., and Fürtig, B. (2021). Switching at the ribosome: riboswitches need rProteins as modulators to regulate translation. *Nat. Commun.* *12*.
<https://doi.org/10.1038/s41467-021-25024-5>.

80. Soulière, M.F., Altman, R.B., Schwarz, V., Haller, A., Blanchard, S.C., and Micura, R. (2013). Tuning a riboswitch response through structural extension of a pseudoknot. *Proc. Natl. Acad. Sci. U. S. A.* *110*. <https://doi.org/10.1073/pnas.1304585110>.

81. Irigoyen, N., Firth, A.E., Jones, J.D., Chung, B.Y.W., Siddell, S.G., and Brierley, I. (2016). High-resolution analysis of coronavirus gene expression by RNA sequencing and ribosome profiling. *PLoS Pathog.* *12*.
<https://doi.org/10.1371/journal.ppat.1005473>.

82. Puray-Chavez, M., Lee, N., Tennen, K., Wang, Y., Vuong, H.R., Liu, Y., Horani, A., Huang, T., Gunsten, S.P., Case, J.B., et al. (2022). The translational landscape of SARS-CoV-2-infected cells reveals suppression of innate immune genes. *mBio* *13*.
<https://doi.org/10.1128/mbio.00815-22>.

83. Kim, D., Kim, S., Park, J., Chang, H.R., Chang, J., Ahn, J., Park, H., Park, J., Son, N., Kang, G., et al. (2021). A high-resolution temporal atlas of the SARS-CoV-2 transcriptome and transcriptome. *Nat. Commun.* *12*. <https://doi.org/10.1038/s41467-021-25361-5>.

84. Cook, G.M., Brown, K., Shang, P., Li, Y., Soday, L., Dinan, A.M., Tumescheit, C., Adrian Mockett, A.P., Fang, Y., Firth, A.E., et al. (2022). Ribosome profiling of porcine reproductive and respiratory syndrome virus reveals novel features of viral gene expression. *Elife* *11*. <https://doi.org/10.7554/ELIFE.75668>.

85. Parmar, S., Bume, D.D., Connelly, C.M., Boer, R.E., Prestwood, P.R., Wang, Z., Labuhn, H., Sinnadurai, K., Feri, A., Ouellet, J., et al. (2024). Mechanistic analysis of riboswitch ligand interactions provides insights into pharmacological control over gene expression. *Nature Communications* *15*. <https://doi.org/10.1038/s41467-024-52235-3>.

86. Wu, M.J., Andreasson, J.O.L., Kladwang, W., Greenleaf, W., and Das, R. (2019). Automated design of diverse stand-alone riboswitches. *ACS Synth. Biol.* *8*, 1838–1846. <https://doi.org/10.1021/acssynbio.9b00142>.

87. Petoukhov, M. V., Konarev, P. V., Kikhney, A.G., and Svergun, D.I. (2007). ATSAS 2.1 – towards automated and web-supported small-angle scattering data analysis. *J Appl. Cryst* *40*, 223–228. <https://doi.org/10.1107/S0021889807002853>

- 1284 88. Konarev, P. V., and Svergun, D.I. (2015). A posteriori determination of the useful data
1285 range for small-angle scattering experiments on dilute monodisperse systems. *IUCrJ*
1286 2, 352–360. <https://doi.org/10.1107/S2052252515005163>.
- 1287 89. Winter, G. (2010). Xia2: An expert system for macromolecular crystallography data
1288 reduction. *J. Appl. Crystallogr.* 43, 186–190.
1289 <https://doi.org/10.1107/S0021889809045701>.
- 1290 90. Kabsch, W. (2010). XDS. *Acta Crystallogr. D Biol. Crystallogr.* 66, 125–132.
1291 <https://doi.org/10.1107/S0907444909047337>.
- 1292 91. Evans, P.R., and Murshudov, G.N. (2013). How good are my data and what is the
1293 resolution? *Acta Crystallogr. D Biol. Crystallogr.* 69, 1204–1214.
1294 <https://doi.org/10.1107/S0907444913000061>.
- 1295 92. Karplus, P.A., and Diederichs, K. (2015). Assessing and maximizing data quality in
1296 macromolecular crystallography. *Curr. Opin. Struct. Biol.* 34, 60–68.
1297 <https://doi.org/10.1016/j.sbi.2015.07.003>.
- 1298 93. McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and
1299 Read, R.J. (2007). Phaser crystallographic software. *J. Appl. Crystallogr.* 40, 658–
1300 674. <https://doi.org/10.1107/S0021889807021206>.
- 1301 94. Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and
1302 development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* 66, 486–501.
1303 <https://doi.org/10.1107/S0907444910007493>.
- 1304 95. Liebschner, D., Afonine, P. V., Baker, M.L., Bunkoczi, G., Chen, V.B., Croll, T.I., Hintze,
1305 B., Hung, L.W., Jain, S., McCoy, A.J., et al. (2019). Macromolecular structure
1306 determination using X-rays, neutrons and electrons: Recent developments in Phenix.
1307 *Acta Crystallogr. D Struct. Biol.* 75, 861–877.
1308 <https://doi.org/10.1107/S2059798319011471>.
- 1309 96. Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J.,
1310 Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: All-atom
1311 structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol.*
1312 *Crystallogr.* 66, 12–21. <https://doi.org/10.1107/S0907444909042073>.
- 1313 97. Schrödinger, LLC. The PyMOL Molecular Graphics System.
- 1314 98. Meng, E.C., Goddard, T.D., Pettersen, E.F., Couch, G.S., Pearson, Z.J., Morris, J.H.,
1315 and Ferrin, T.E. (2023). UCSF ChimeraX: Tools for structure building and analysis.
1316 *Protein Science* 32. <https://doi.org/10.1002/pro.4792>.
- 1317 99. Li, S., Olson, W.K., and Lu, X.J. (2019). Web 3DNA 2.0 for the analysis, visualization,
1318 and modeling of 3D nucleic acid structures. *Nucleic Acids Res.* 47, W26–W34.
1319 <https://doi.org/10.1093/nar/gkz394>.
- 1320 100. Blanchet, C.E., Spilotros, A., Schwemmer, F., Graewert, M.A., Kikhney, A., Jeffries,
1321 C.M., Franke, D., Mark, D., Zengerle, R., Cipriani, F., et al. (2015). Versatile sample
1322 environments and automation for biological solution X-ray scattering experiments at
1323 the P12 beamline (PETRA III, DESY). *J. Appl. Crystallogr.* 48, 431–443.
1324 <https://doi.org/10.1107/S160057671500254X>.

- 1325 101. Panjkovich, A., and Svergun, D.I. (2018). CHROMIXS: Automatic and interactive
1326 analysis of chromatography-coupled small-angle X-ray scattering data. *Bioinformatics*
1327 *34*, 1944–1946. <https://doi.org/10.1093/bioinformatics/btx846>.
- 1328 102. Franke, D., Kikhney, A.G., and Svergun, D.I. (2012). Automated acquisition and
1329 analysis of small angle X-ray scattering data. *Nucl. Instrum. Methods Phys. Res. A*
1330 *689*, 52–59. <https://doi.org/10.1016/J.NIMA.2012.06.008>.
- 1331 103. Franke, D., Jeffries, C.M., and Svergun, D.I. (2015). Correlation Map, a goodness-of-
1332 fit test for one-dimensional X-ray scattering spectra. *Nat. Methods* *12*, 419–422.
1333 <https://doi.org/10.1038/nmeth.3358>.
- 1334 104. Manalastas-Cantos, K., Konarev, P. V., Hajizadeh, N.R., Kikhney, A.G., Petoukhov, M.
1335 V., Molodenskiy, D.S., Panjkovich, A., Mertens, H.D.T., Gruzinov, A., Borges, C., et al.
1336 (2021). ATSAS 3.0: Expanded functionality and new tools for small-angle scattering
1337 data analysis. *J. Appl. Crystallogr.* *54*, 343–355.
1338 <https://doi.org/10.1107/S1600576720013412>.
- 1339 105. Petoukhov, M. V., Franke, D., Shkumatov, A. V., Tria, G., Kikhney, A.G., Gajda, M.,
1340 Gorba, C., Mertens, H.D.T., Konarev, P. V., and Svergun, D.I. (2012). New
1341 developments in the ATSAS program package for small-angle scattering data
1342 analysis. *J. Appl. Crystallogr.* *45*, 342–350.
1343 <https://doi.org/10.1107/S0021889812007662>.
- 1344 106. Rambo, R.P., and Tainer, J.A. (2013). Accurate assessment of mass, models and
1345 resolution by small-angle scattering. *Nature* *496*, 477–481.
1346 <https://doi.org/10.1038/nature12070>.
- 1347 107. Fischer, H., De Oliveira Neto, M., Napolitano, H.B., Polikarpov, I., and Craievich, A.F.
1348 (2010). Determination of the molecular weight of proteins in solution from a single
1349 small-angle X-ray scattering measurement on a relative scale. *J. Appl. Crystallogr.* *43*,
1350 101–109. <https://doi.org/10.1107/S0021889809043076>.
- 1351 108. Hajizadeh, N.R., Franke, D., Jeffries, C.M., and Svergun, D.I. (2018). Consensus
1352 Bayesian assessment of protein molecular mass from solution X-ray scattering data.
1353 *Sci. Rep.* *8*. <https://doi.org/10.1038/s41598-018-25355-2>.
- 1354 109. Svergun, D.I. (1992). Determination of the regularization parameter in indirect-
1355 transform methods using perceptual criteria. *J. Appl. Cryst* *25*, 495–503.
1356 <https://doi.org/10.1107/S0021889892001663>.
- 1357 110. Svergun, D.I. (1999). Restoring low resolution structure of biological macromolecules
1358 from solution scattering using simulated annealing. *Biophys. J.* *76*, 2879–2886.
1359 [https://doi.org/10.1016/S0006-3495\(99\)77443-6](https://doi.org/10.1016/S0006-3495(99)77443-6).
- 1360 111. Whitten, A.E., Cai, S., and Trehwella, J. (2008). MULCh: Modules for the analysis of
1361 small-angle neutron contrast variation data from biomolecular assemblies. *J. Appl.*
1362 *Crystallogr.* *41*, 222–226. <https://doi.org/10.1107/S0021889807055136>.
- 1363 112. Trehwella, J., Duff, A.P., Durand, D., Gabel, F., Guss, J.M., Hendrickson, W.A., Hura,
1364 G.L., Jacques, D.A., Kirby, N.M., Kwan, A.H., et al. (2017). 2017 publication guidelines
1365 for structural modelling of small-angle scattering data from biomolecules in solution:
1366 An update. *Acta Crystallogr. D Struct. Biol.* *73*, 710–728.
1367 <https://doi.org/10.1107/S2059798317011597>.

1368 113. Panjkovich, A., and Svergun, D.I. (2016). Deciphering conformational transitions of
1369 proteins by small angle X-ray scattering and normal mode analysis. *Physical*
1370 *Chemistry Chemical Physics* *18*, 5707–5719. <https://doi.org/10.1039/c5cp04540a>.

1371 114. Svergun, D., Barberato, C., and Koch, M.H.J. (1995). CRY SOL-a Program to Evaluate
1372 X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J*
1373 *Appl. Cryst* *28*, 768–773. <https://doi.org/10.1107/S0021889895007047>

1374 115. Kikhney, A.G., Borges, C.R., Molodenskiy, D.S., Jeffries, C.M., and Svergun, D.I.
1375 (2020). SASBDB: Towards an automatically curated and validated repository for
1376 biological scattering data. *Protein Science* *29*, 66–75.
1377 <https://doi.org/10.1002/pro.3731>.

1378 116. Ingargiola, A., Lerner, E., Chung, S.Y., Weiss, S., and Michalet, X. (2016).
1379 FRET Bursts: An open source toolkit for analysis of freely-diffusing Single-molecule
1380 FRET. *PLoS One* *11*. <https://doi.org/10.1371/journal.pone.0160716>.

1381 117. Abdelhamid, M.A.S., Rhind-Tutt, A. V., Ambrose, B., and Craggs, T.D. (2021). Making
1382 precise and accurate single-molecule FRET measurements using the open-source
1383 smfBox. *Journal of Visualized Experiments* *2021*. <https://doi.org/10.3791/62378>.

1384 118. Hellenkamp, B., Schmid, S., Doroshenko, O., Opanasyuk, O., Kühnemuth, R., Rezaei
1385 Adariani, S., Ambrose, B., Aznauryan, M., Barth, A., Birkedal, V., et al. (2018).
1386 Precision and accuracy of single-molecule FRET measurements—a multi-laboratory
1387 benchmark study. *Nat. Methods* *15*, 669–676. [https://doi.org/10.1038/s41592-018-](https://doi.org/10.1038/s41592-018-0085-0)
1388 *0085-0*.

1389 119. Kalinin, S., Peulen, T., Sindbert, S., Rothwell, P.J., Berger, S., Restle, T., Goody, R.S.,
1390 Gohlke, H., and Seidel, C.A.M. (2012). A toolkit and benchmark study for FRET-
1391 restrained high-precision structural modeling. *Nat. Methods* *9*, 1218–1225.
1392 <https://doi.org/10.1038/nmeth.2222>.

1393 120. Loughran, G., Fedorova, A.D., Khan, Y.A., Atkins, J.F., and Baranov, P. V. (2022). Lack
1394 of evidence for ribosomal frameshifting in ATP7B mRNA decoding. *Mol. Cell* *82*, 3745-
1395 3749.e2. <https://doi.org/10.1016/j.molcel.2022.08.024>.

1396 121. Wood, G., Lastovka, F., Murza, D., Brember, M.P., McGovern, N., and Chung, B.Y.W.
1397 (2025). An RNA-Based Dual-Fluorescence Reporter System Reveals Cell Type-
1398 Specific and Temporal Dynamics of –1 Programmed Ribosomal Frameshifting. *J. Mol.*
1399 *Biol.* <https://doi.org/10.1016/j.jmb.2025.169371>.

1400

Supplementary Material

Movie S1. An annotated video highlighting key structural features of TMEV 2A protein bound to pseudoknot RNA. <https://www.youtube.com/watch?v=I3ji-RdrPP8> (Related to **Figure 2**).

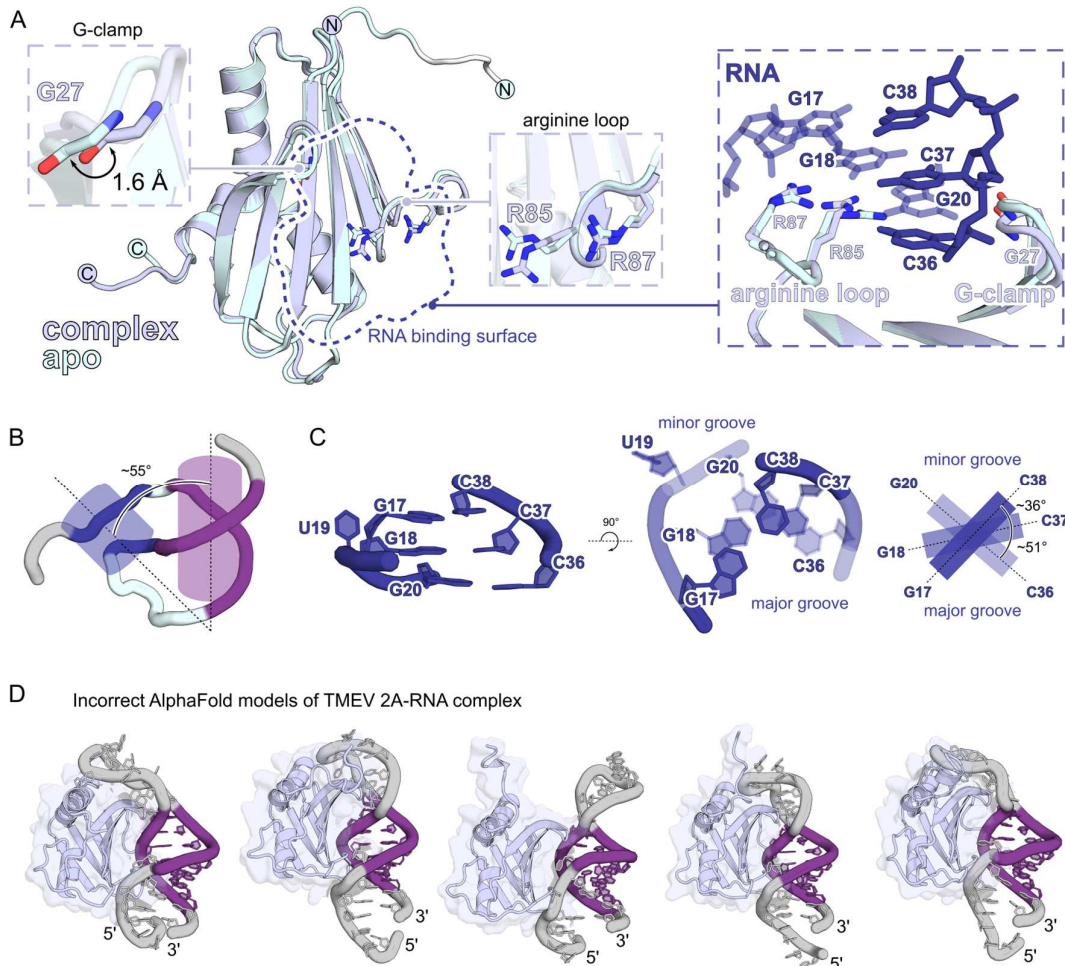


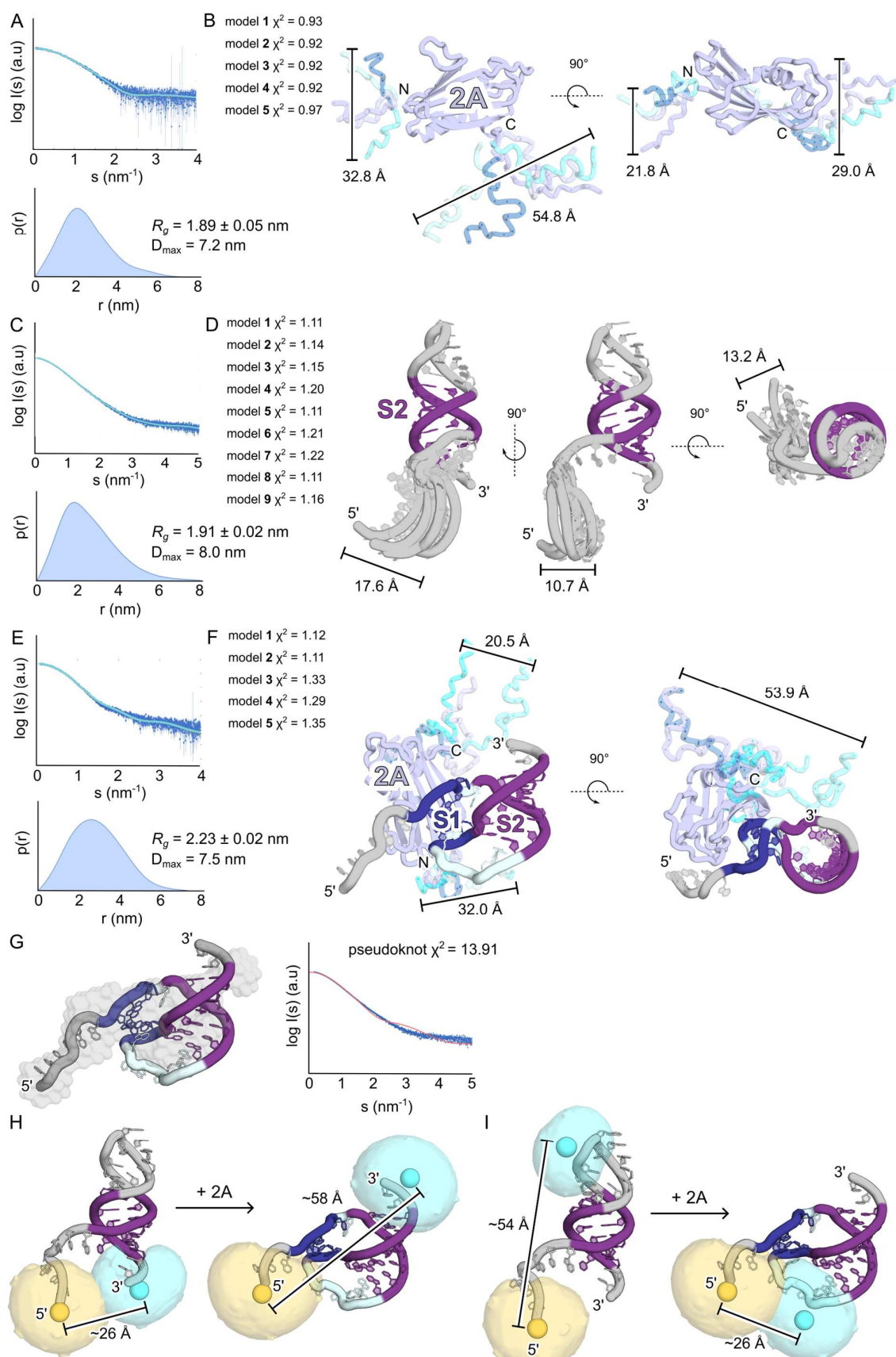
Figure S1. Comparison of the *apo* and *holo* TMEV 2A protein structures. Related to **Figures 1 and 2**, and **Table 1**.

A) Alignment of the *apo* (PDB 7NBV [1]) and *holo* (this study, PDB 9RVP) 2A protein structures reveals modest conformational plasticity of the ‘G-clamp’ (G27) and arginine loop (R85/R87) involved in RNA recognition. (*Insets*) Close-up of the ‘G-clamp’ and arginine loop shown as sticks.

B) Overview of the RNA pseudoknot showing the non-coaxial stacking of the two stems. The helical axes are represented by dotted black line and cartoon cylinders. The angle between the helical axis of S1 and S2 is ~55°.

C) Details of the non-consecutive guanine nucleotides in S1 that creates an unusual helical geometry. (*Left*) Side view showing the planarity of the three G-C base pairs. (*Middle*) Top view showing the effect of the flipped base U19 to the helical twist of S1. (*Right*) Schematic of S1 along the helical axes with helical twist angles indicated. The first and second base-pairs (G17-C38 and G18-C37) have a twist of ~36° whereas the twist of the second and third base-pair (G18-C37 and G20-C36) is ~51°. Helical twist angles analysed using the web 3DNA server [2].

D) AlphaFold3 [3] predictions of the 2A-RNA complex. An incorrect side-by-side arrangement is predicted in which 2A interacts with S2.



1422 **Figure S2 (previous page).** Aligned atomistic models consistent with solution SAXS data for 2A protein,
 1423 RNA, and 2A-RNA complex, demonstrating principal modes of flexibility. Related to **Figure 3** and **Table**
 1424 **2.**

1425 A) Plots showing alignment between SAXS data for the 2A protein (dark blue circles) and calculated
 1426 data from structural models (cyan lines), and the pair-distance distribution $p(r)$ allowing the
 1427 determination of R_g and D_{max} . Mean R_g values in nm are reported $\pm$ SD.

1428 B) Flexible N- and C-termini of 2A (shades of blue) were modelled in CORAL [4]. Five superposed
 1429 atomistic models are presented that fit the SAXS data, as assessed by reduced χ^2 test. Displacements
 1430 between equivalent termini are indicated in angstroms (Å).

1431 C) Plots as in A) for RNA alone. Mean R_g values in nm are reported $\pm$ SD.

1432 D) Normal mode analysis using SREFLEX [5] illustrates conformational flexibility of the 5' extension.
 1433 Nine superposed atomistic models are presented that fit the SAXS data, as assessed by reduced χ^2
 1434 test. Displacements between equivalent 5' nucleotides are indicated in angstroms (Å).

1435 E) Plots as in A) for 2A-RNA complex. Mean R_g values in nm are reported $\pm$ SD.

1436 F) Flexible N- and C-termini (shades of blue) of 2A were modelled in CORAL [4]. Five superposed
 1437 atomistic models are presented that fit the SAXS data, as assessed by reduced χ^2 test. Displacements
 1438 between equivalent termini are indicated in angstroms (Å).

1439 G) The pseudoknot conformation observed in the crystal structure does not explain the RNA-only
 1440 scattering data, as assessed by reduced χ^2 test. (*Left*) Pseudoknot model from crystal structure docked
 1441 into SEC-SAXS envelope from RNA-only scattering data. (*Right*) Plots showing alignment between
 1442 SAXS data for the RNA alone (dark blue circles) and calculated data from pseudoknot model (red line).

1443 H, I) Accessible volume (AV) simulations calculated using FRET-restrained positioning and screening
 1444 (FPS) tool [6] of RNA reporter constructs used in smFRET experiments. AV for Cy5 (yellow) and Cy3
 1445 (cyan) are represented as transparent surface and mean dye points as spheres. Distances from mean
 1446 dye points were measured in PyMol [7]. H) end reporter. I) internal reporter.

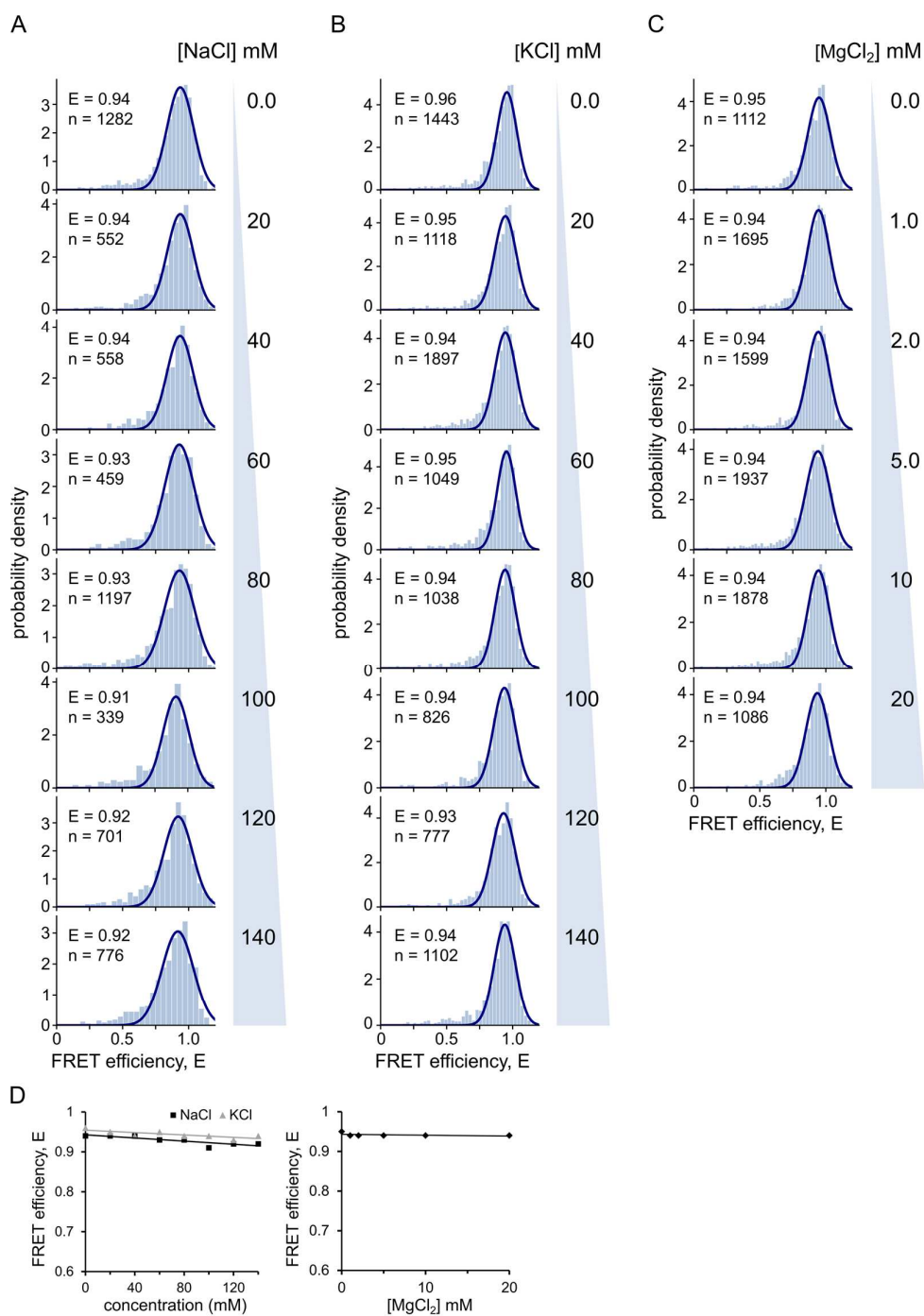


Figure S3. Ionic strength and identity have no effect on the conformation of the RNA. Related to Figure 3.

A-C) FRET efficiency (E) histograms of the 'end reporter' TMEV RNA fitted to a single Gaussian distribution. Corrected FRET efficiencies (E) and number of molecules (n) reported in each histogram. Titrations of A) NaCl B) KCl C) MgCl₂.

D) Summary of the 2D histograms with the corrected FRET efficiency plotted against ionic concentration. NaCl is represented by black squares and KCl by grey triangles.

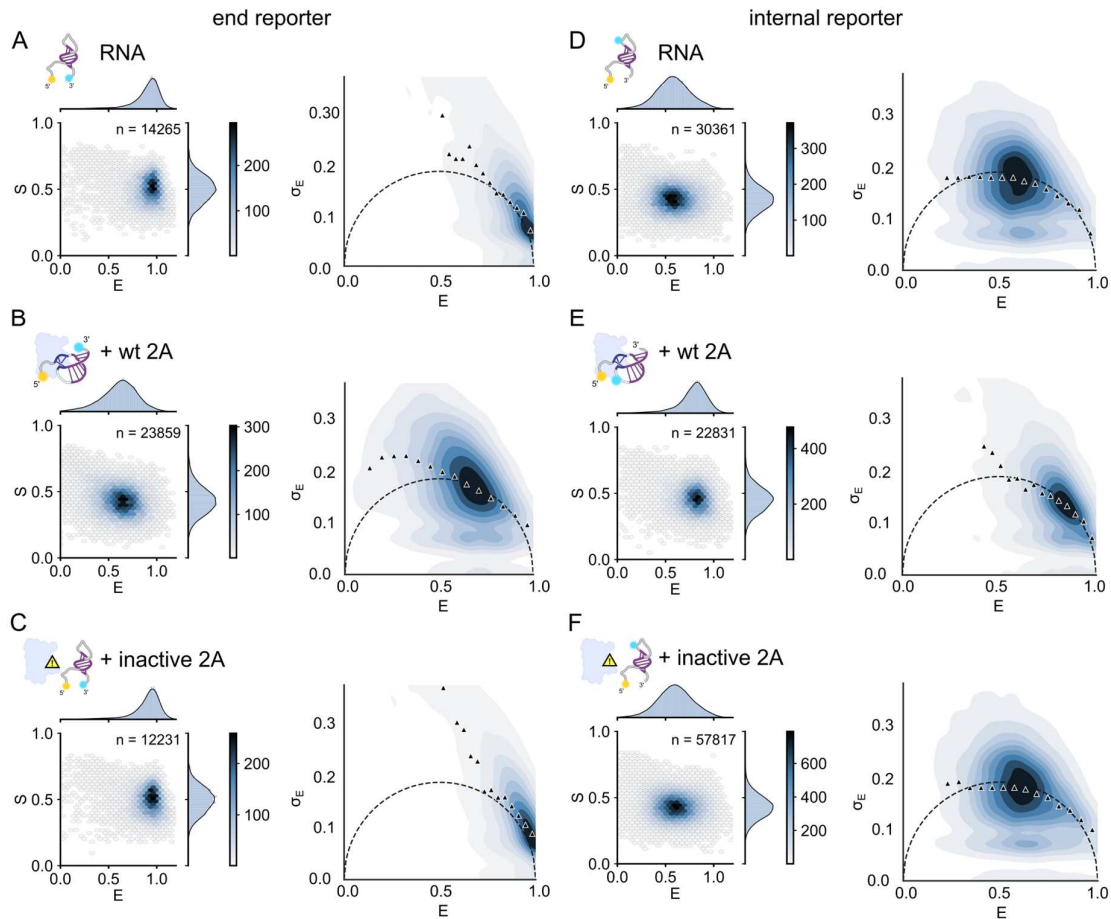


Figure S4. The RNA does not spontaneously convert between multiple stable conformations. Related to **Figure 3**.

(Left) 2D ALEX histograms of stoichiometry (S) against FRET efficiency (E) after filtering out the donor and acceptor only populations. The number of molecules in each dataset (n) is indicated. Data was collected from three independent experiments and combined. (Right) Corresponding burst variance analysis (BVA) contour plots assessing conformational dynamics of the RNA stimulatory element. Bursts were divided into sub-bursts of seven photons and the standard deviation calculated for each sub-burst population, indicated by black triangles. A standard deviation of sub-bursts greater than the theoretical shot-noise-limited variance (dashed black curve) suggests dynamics.

A-C) end reporter RNA (5' Cy5 and 3' Cy3). D-F) internal reporter RNA (5' Cy5 and Cy3 positioned on U31). A,D) RNA reporters in isolation. B,E) in the presence of 0.5 μ M wild type 2A. C,F) in the presence of 0.5 μ M inactive 2A_{R85A/R87A}.

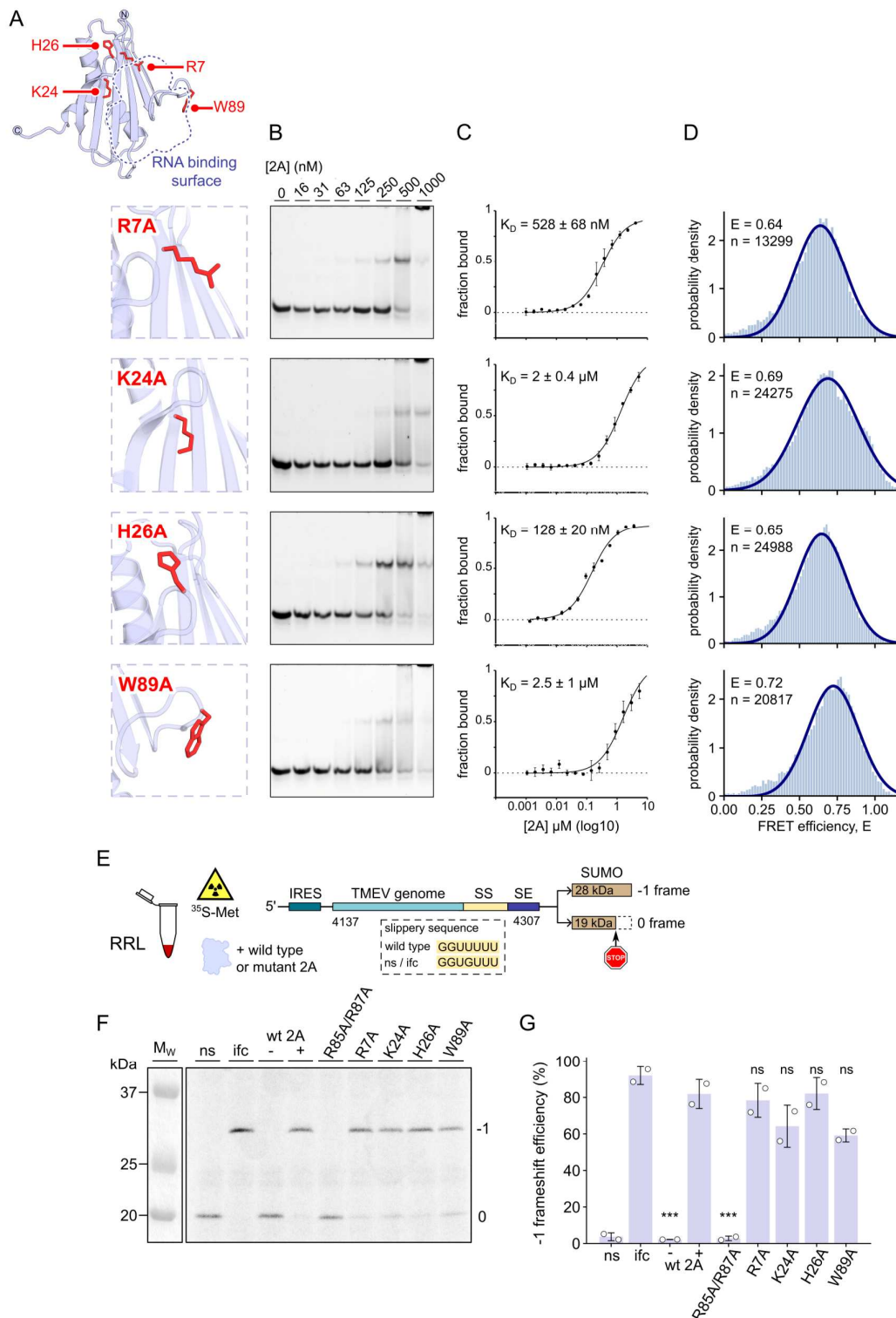


Figure S5. Mutagenesis of protein residues at the RNA binding interface has little effect on RNA recognition. Related to **Figures 4 and 5**.

A) Mutagenesis of residues in TMEV 2A that are involved in RNA binding. The locations of the mutations R7A, K24A, H26A and W89A are shown as red sticks. The RNA binding surface is represented as a blue dashed line.

1474 B) EMSA analysis of the effects of mutant 2A proteins on RNA stimulatory element binding. Experiments
 1475 were performed with 50 nM wild type RNA labelled with Cy5 at the 5' end. 2A concentrations were
 1476 assessed between 0-1.0 μ M. Following non-denaturing gel electrophoresis the gels were imaged with
 1477 a Typhoon-5 scanner.

1478 C) MST binding curves and approximate K_D values using the same mutant 2A proteins. The K_D fit model
 1479 was used to describe a 1:1 binding stoichiometry. Error bars indicate standard deviation from the mean
 1480 of three biological repeats.

1481 D) Combined FRET efficiency (E) histograms of end reporter wild type RNA with mutant 2A proteins
 1482 from three biological repeats. The distribution is fitted to a Gaussian distribution where $R^2 = 0.98$.
 1483 Average FRET efficiency (E) and number of molecules (n) is reported on each histogram.

1484 E) Schematic diagram of *in vitro* translation assay used to assess the effects of 2A mutations on
 1485 frameshifting efficiency. Reporter constructs contained nucleotides 4137-4307 of the TMEV genome
 1486 encompassing 35 codons of the viral genome prior to the viral frameshift site (coloured in blue). The
 1487 slippery sequence (SS, yellow) and stimulatory element (SE, dark blue) are highlighted. SUMO protein
 1488 (brown) was used as the product of translation in either the 0 or -1 frames. If elongation continues in
 1489 the 0 frame the resulting product is ~19 kDa. Shifting into the -1 frame results in a ~28 kDa product.
 1490 The non-slip (ns) control mRNA contains a mutated slippery sequence to prevent shifting into the -1
 1491 frame. The in-frame control (ifc) mRNA introduces another nucleotide after the stimulatory element, to
 1492 simulate the effect of 100% -1 frameshifting. Translation reactions were carried out in rabbit reticulocyte
 1493 lysate (RRL) in the presence of 35 S-Met, analysed by SDS-PAGE and visualised by autoradiography.

1494 F) SDS-PAGE and autoradiography analysis of 2A mutations on frameshifting efficiency *in vitro*.

1495 G) Densitometric quantification of the -1 frameshift efficiencies, corrected for methionine content. Data
 1496 are represented as the mean $\pm$ SD (n = 2). A one-way ANOVA with Tukey HSD test was used to compare
 1497 2A mutants to wild type RNA + 2A. ns, not significant, *** $p \leq 0.001$.

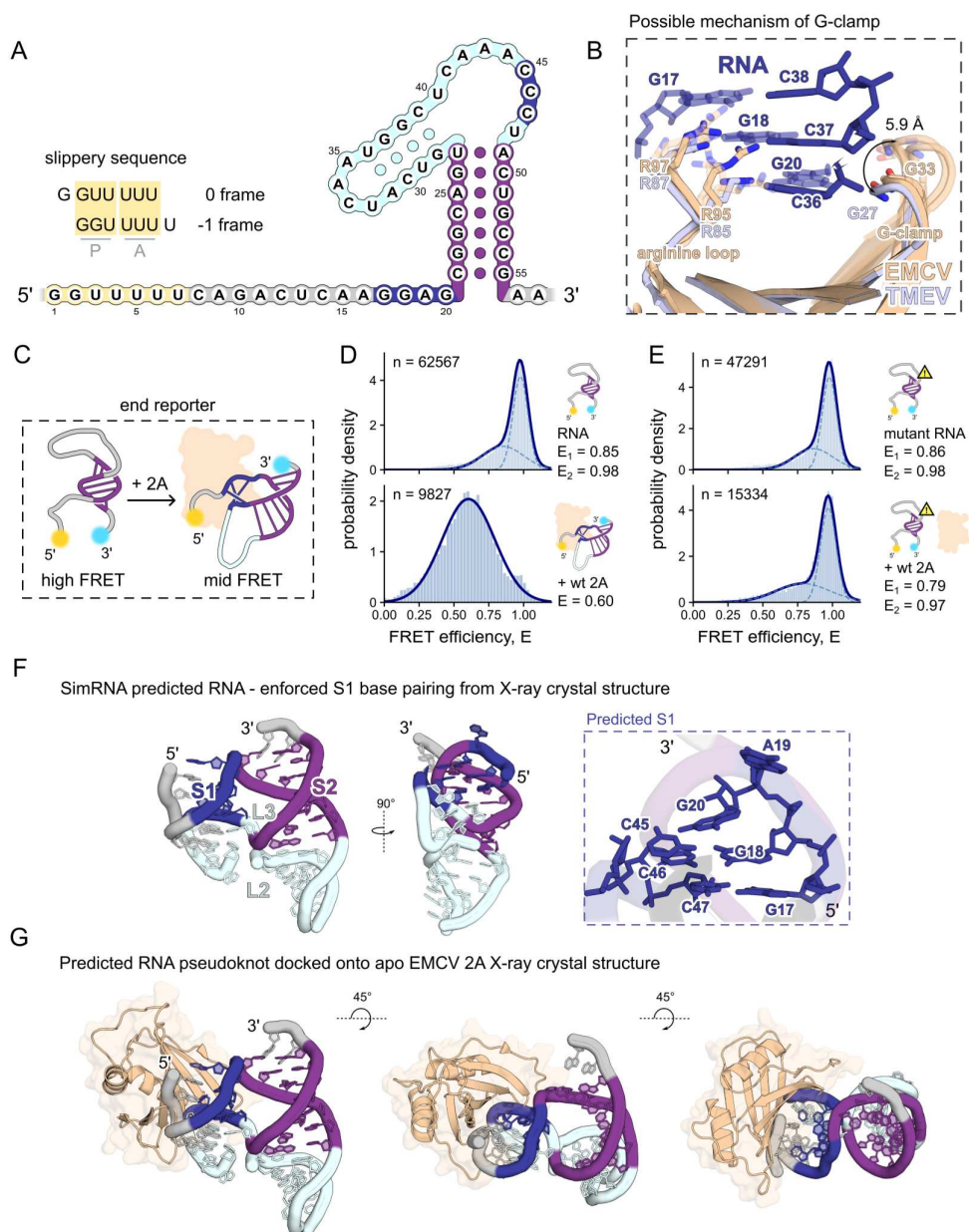


Figure S6. Proposed riboswitch mechanism in EMCV. Related to **Figures 1, 2 and 3.**

A) Schematic of the EMCV (47 nt) minimal RNA stimulatory element represented as a stem-loop. Slippery sequence is coloured in yellow, nucleotides not involved in base pairing are coloured in grey. Highly conserved GGxG and CCC motifs that comprise S1 coloured in dark blue, S2 coloured in purple and loops in cyan.

B) Alignment of *apo* EMCV crystal structure (wheat) (PDB 7BNY [8]) with TMEV 2A-RNA pseudoknot structure – this study (2A, light blue; RNA S1, dark blue). The RNA binding surface is likely to be conserved across the cardioviruses. NCS-related chains in the asymmetric unit of the *apo* EMCV 2A crystal structure (PDB 7BNY [8]) reveals a possible mechanism for the 'G-clamp' which demonstrates a pincer motion.

C) Schematic diagrams showing smFRET EMCV RNA end reporter construct design using the Cy3 (cyan) and Cy5 (yellow) FRET pair.

1511 D) Corrected FRET efficiency (E) histograms, data combined from three biological repeats. Corrected
1512 average FRET efficiencies (E) and number of molecules (n) are reported in each histogram. Gaussian
1513 fits to our data are shown as dashed curves for two Gaussians and a solid blue line for a single Gaussian
1514 fit. $R^2 = 0.99$ for all Gaussian fits. Showing data for (*upper*) RNA in isolation, (*lower*) RNA in the presence
1515 of 0.5 μ M wild type EMCV 2A.

1516 E) As in D), but using EMCV C47G mutant RNA, predicted to prevent the formation of S1 (equivalent
1517 to TMEV C36G).

1518 F) SimRNA [9] predicted model of the EMCV RNA pseudoknot with enforced S1 GGxG base pairing in
1519 two orthogonal views. (*Inset*) Close-up view of predicted S1 with flipped out base A19 in EMCV.

1520 G) Predicted model of the EMCV 2A-RNA pseudoknot complex. Using the TMEV 2A-RNA pseudoknot
1521 complex as a reference, the *apo* EMCV 2A structure (PDB 7BNY) was aligned and docked onto the
1522 SimRNA [9] predicted EMCV RNA pseudoknot, suggesting a possible conserved mechanism of 2A-
1523 induced pseudoknot formation.

1524

1525

Table S1. DNA oligonucleotide sequences. Related to **STAR Methods**.

Primer pair	Sequence	Source
1	Primers for subcloning TMEV 2A protein (wt and inactive 2A _{R85A/R87A}) (F) 5' - AATTCATATGAATCCCGCTTCTCTCTACCG - 3' (R) 3' - AATTGGATCCCTAGCCTGGGTTTCATTTCTACATCATG - 5'	Integrated DNA Technologies (IDT)
2	Mutagenesis PCR primers – 2A _{R7A} (F) 5' - CCCGCTTCTCTCTACGCCATTGATCTTTTCAT - 3' (R) 3' - ATGAAAAGATCAATGGCGTAGAGAGAAGCGGG - 5'	Integrated DNA Technologies (IDT)
3	Mutagenesis PCR primers – 2A _{K24A} (F) 5' - ATAACCTTTTACTACGCGGTCCACGGACGTCC - 3' (R) 3' - GGACGTCCGTGGACCGCGTAGTCAAAAAGTTAT - 5'	Integrated DNA Technologies (IDT)
4	Mutagenesis PCR primers – 2A _{H26A} (F) 5' - TTTGACTACAAGGTCGCCGACGTCCTGTGCT - 3' (R) 3' - AGCACAGGACGTCCGCGACCTTGTAAGTCAAA - 5'	Integrated DNA Technologies (IDT)
5	Mutagenesis PCR primers – 2A _{W89A} (F) 5' - GACGCTACCGCTCCGCGAAGAAGCCCATCCAC - 3' (R) 3' - GTGGATGGGCTTCTTCGCGGAGCGGTAGCGTC - 5'	Integrated DNA Technologies (IDT)
6	Primers for cloning wild type and test sequences into pSGDlucV3.0 (F) 5' - CAGACTCGAGTCTCGAGTACACCACTTTCCG - 3' (R) 3' - ACATAGATCTTTGTTTTGAATTCCTGCCTGG - 5'	Merck
7	Primers for cloning the in-frame control (ifc) into pSGDlucV3.0 (F) 5' - CAGACTCGAGTCTCGAGTACACCACTTTCCG - 3' (R) 3' - ACATAGATCTGTTTTGAATTCCTGCCTGG - 5'	Merck
8	Primers for cloning - 1 bp stem 2 into pSGDlucV3.0 (F) 5' - CAGACTCGAGTCTCGAGTACACCACTTTCCG - 3' (R) 3' - ACATAGATCTTTGTTTTTTGAATTCCTGCCTGG - 5'	Merck
9	Primers for cloning - 2 bp stem 2 into pSGDlucV3.0 (F) 5' - CAGACTCGAGTCTCGAGTACACCACTTTCCG - 3' (R) 3' - ACATAGATCTTTGTTTTTTTTGAATTCCTGCCTGG - 5'	Merck
10	Primers for cloning TMEV 2A into pCEP4 for dual luciferase assays (F) 5' - CATATGCTCGAGCCGCCACCATGAATCCCGCTTCTCTCTAC - 3' (R) 3' - TTAGATCTGGATCCCTAGCCTGGGTTTCATTTCTAC - 5'	Merck

1527 **Table S2.** RNA oligonucleotide sequences. Related to **STAR Methods**.

Oligo number	Sequence	Source
1	RNA used in crystallisation of 2A-RNA complex 5' - CAAGGUGCGGUGCUAACUAAAUCCCUAGCACCCC - 3'	Integrated DNA Technologies (IDT)
2	Cy3/Cy5 'end reporter' RNA used in smFRET experiments 5' - Cy5_CAAGGUGCGGUGCUAACUAAAUCCCUAGCACCCC_Cy3 - 3'	Integrated DNA Technologies (IDT)
3	Cy3/Cy5 'internal reporter' RNA used in smFRET experiments 5' - Cy5_CAAGGUGCGGUGCUAAC(U-Cy3)AAAUCCCUAGCACCCC - 3'	Integrated DNA Technologies (IDT)
4	Stem 1 swap used in EMSA and MST 5' - CAACCUCCGUGCUAACUAAAUGGGUAGCACCCC - 3'	Merck
5	AU stem 1 used in EMSA and MST 5' - CAAAUAACGGUGCUAACUAAAUUUUUAGCACCCC - 3'	Merck
6	UA stem 1 used in EMSA and MST 5' - CAAUUUUCCGUGCUAACUAAAUAAUAGCACCCC - 3'	Merck
7	Δ U19 stem 1 used in EMSA and MST 5' - CAAGG_GCGGUGCUAACUAAAUCCCUAGCACCCC - 3'	Merck
8	Stem 2 swap used in EMSA and MST 5' - CAAGGUGCCCACGAUACUAAAUCCCAUCGUGGCC - 3'	Integrated DNA Technologies (IDT)
9	+ 1 bp stem 2 used in EMSA and MST 5' - CAAGGUGGGGUGCUUACUAAAUCCCAAGCACCCC- 3'	Integrated DNA Technologies (IDT)
10	-1 bp stem 2 used in EMSA and MST 5' - CAAGGUGCGGUGCUACUAAAUCCCAAGCACCCC- 3'	Integrated DNA Technologies (IDT)
11	-2 bp stem 2 used in EMSA and MST 5' - CAAGGUGCGGUGCACUAAAUCCCGCACCCC- 3'	Integrated DNA Technologies (IDT)
12	GGG mutant used in EMSA and MST 5' - CAAGGUGCGGUGCUAACUAAAUGGGUAGCACCCC - 3'	Merck
13	C38G mutant RNA used in EMSA and MST 5' - CAAGGUGCGGUGCUAACUAAAUCCGUAGCACCCC - 3'	Merck
14	C37G mutant RNA used in EMSA and MST 5' - CAAGGUGCGGUGCUAACUAAAUCGCUAGCACCCC - 3'	Merck
15	C36G mutant RNA used in EMSA and MST 5' - CAAGGUGCGGUGCUAACUAAAUGCCUAGCACCCC - 3'	Merck
16	Cy3/Cy5 GGG mutant RNA used in smFRET experiments 5' - Cy5_CAAGGUGCGGUGCUAACUAAAUGGGUAGCACCCC_Cy3 - 3'	Integrated DNA Technologies (IDT)
17	Cy3/Cy5 C38G mutant RNA used in smFRET experiments 5' - Cy5_CAAGGUGCGGUGCUAACUAAAUCCGUAGCACCCC _Cy3 - 3'	Integrated DNA Technologies (IDT)
18	Cy3/Cy5 C37G mutant RNA used in smFRET experiments 5' - Cy5_CAAGGUGCGGUGCUAACUAAAUCGCUAGCACCCC_Cy3 - 3'	Integrated DNA Technologies (IDT)
19	Cy3/Cy5 C36G mutant RNA used in smFRET experiments 5' - Cy5_CAAGGUGCGGUGCUAACUAAAUGCCUAGCACCCC_Cy3 - 3'	Integrated DNA Technologies (IDT)

20	Cy3/Cy5 EMCV wt RNA used in smFRET experiments 5' - Cy5_ACUCAAGGAGCGGCAGUGUCAUCAAAUGGCUCAAACCC UACUGCCGAA_Cy3 - 3'	Integrated DNA Technologies (IDT)
21	Cy3/Cy5 C47G mutant EMCV RNA used in smFRET experiments 5' - Cy5_ACUCAAGGAGCGGCAGUGUCAUCAAAUGGCUCAAACCG UACUGCCGAA_Cy3 - 3'	Integrated DNA Technologies (IDT)

1528

1529 **Table S3.** smFRET Gaussian fits parameters. Related to **Figure 3**.

Sample	FRET E ₁	FWHM ^a Gaussian ₁	FRET E ₂	FWHM ^a Gaussian ₂
End reporter RNA	0.85 ± 0.03	0.29 ± 0.03	0.96 ± 0.01	0.14 ± 0.01
End reporter RNA + 2A	0.63 ± 0.01	0.40 ± 0.01	n/a	n/a
End reporter RNA + inactive 2A	0.86 ± 0.02	0.31 ± 0.03	0.96 ± 0.00	0.15 ± 0.01
Internal reporter RNA	0.58 ± 0.00	0.40 ± 0.01	n/a	n/a
Internal reporter RNA + 2A	0.67 ± 0.02	0.47 ± 0.04	0.83 ± 0.01	0.23 ± 0.01
Internal reporter RNA + inactive 2A	0.60 ± 0.01	0.41 ± 0.01	n/a	n/a

1530 ^aFWHM: full width half maxima. Data is presented as the mean ± standard deviation of three
1531 independent replicates.

Supplementary References

1. Hill, C.H., Cook, G.M., Naphthine, S., Kibe, A., Brown, K., Caliskan, N., Firth, A.E., Graham, S.C., and Brierley, I. (2021). Investigating molecular mechanisms of 2A-stimulated ribosomal pausing and frameshifting in Theilovirus. *Nucleic Acids Res.* **49**, 11938–11958. <https://doi.org/10.1093/nar/gkab969>.
2. Li, S., Olson, W.K., and Lu, X.J. (2019). Web 3DNA 2.0 for the analysis, visualization, and modeling of 3D nucleic acid structures. *Nucleic Acids Res.* **47**, W26–W34. <https://doi.org/10.1093/nar/gkz394>.
3. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A.J., Bambrick, J., et al. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500. <https://doi.org/10.1038/s41586-024-07487-w>.
4. Petoukhov, M. V., Franke, D., Shkumatov, A. V., Tria, G., Kikhney, A.G., Gajda, M., Gorba, C., Mertens, H.D.T., Konarev, P. V., and Svergun, D.I. (2012). New developments in the ATSAS program package for small-angle scattering data analysis. *J. Appl. Crystallogr.* **45**, 342–350. <https://doi.org/10.1107/S0021889812007662>.
5. Panjkovich, A., and Svergun, D.I. (2016). Deciphering conformational transitions of proteins by small angle X-ray scattering and normal mode analysis. *Physical Chemistry Chemical Physics* **18**, 5707–5719. <https://doi.org/10.1039/c5cp04540a>.
6. Kalinin, S., Peulen, T., Sindbert, S., Rothwell, P.J., Berger, S., Restle, T., Goody, R.S., Gohlke, H., and Seidel, C.A.M. (2012). A toolkit and benchmark study for FRET-restrained high-precision structural modeling. *Nat. Methods* **9**, 1218–1225. <https://doi.org/10.1038/nmeth.2222>.
7. Schrödinger, LLC. The PyMOL Molecular Graphics System.
8. Hill, C.H., Pekarek, L., Naphthine, S., Kibe, A., Firth, A.E., Graham, S.C., Caliskan, N., and Brierley, I. (2021). Structural and molecular basis for Cardiovirus 2A protein as a viral gene expression switch. *Nat. Commun.* **12**. <https://doi.org/10.1038/s41467-021-27400-7>.
9. Moafinejad, S.N., De Aquino, B.R.H., Boniecki, M.J., Pandaranadar Jeyeram, I.P.N., Nikolaev, G., Magnus, M., Farsani, M.A., Badepally, N.G., Wirecki, T.K., Stefaniak, F., et al. (2024). SimRNAweb v2.0: a web server for RNA folding simulations and 3D structure modeling, with optional restraints and enhanced analysis of folding trajectories. *Nucleic Acids Res.* **52**, W368–W373. <https://doi.org/10.1093/nar/gkae356>.