Structural studies of Cardiovirus 2A protein reveal the molecular basis for RNA recognition and translational control

Chris H. Hill*[†1,3], Sawsan Napthine[†1], Lukas Pekarek[†2], Anuja Kibe[2], Andrew E. Firth[1], Stephen C. Graham*[1], Neva Caliskan*[2,4] and Ian Brierley*[1]

[†] authors contributed equally to this work

* corresponding authors

[1] Division of Virology, Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge, UK. CB2 1QP

[2] Helmholtz Institute for RNA-based Infection Research (HIRI), Josef-Schneider-Straße 2/D15, 97080 Würzburg, Germany

[3] MRC Laboratory of Molecular Biology, Cambridge Biomedical Campus, Francis Crick Ave, Cambridge, UK. CB2 0QH

[4] Medical Faculty, Julius-Maximilians University Würzburg, 97074, Würzburg, Germany

Abstract

Encephalomyocarditis virus 2A protein is a multi-functional virulence factor essential for efficient virus replication with roles in stimulating programmed −1 ribosomal frameshifting (PRF), inhibiting cap-dependent translational initiation, interfering with nuclear import and export and preventing apoptosis of infected cells. The mechanistic basis for many of these activities is unclear and a lack of structural data has hampered our understanding. Here we present the X-ray crystal structure of 2A, revealing a novel "beta-shell" fold. We show that 2A selectively binds to and stabilises a specific conformation of the stimulatory RNA element in the viral genome that directs PRF at the 2A/2B* junction. We dissect the folding energy landscape of this stimulatory RNA element, revealing multiple conformers, and measure changes in unfolding pathways arising from mutation and 2A binding. Furthermore, we demonstrate a strong interaction between 2A and the small ribosomal subunit and present a high-resolution cryo-EM structure of 2A bound to initiated 70S ribosomes. In this complex, three copies of 2A bind directly to 16S ribosomal RNA at the factor binding site, where they may compete for binding with initiation and elongation factors. Together, these results provide an integrated view of the structural basis for RNA recognition by 2A, expand our understanding of PRF, and provide unexpected insights into how a multifunctional viral protein may shut down translation during virus infection.

Introduction

Encephalomyocarditis virus (EMCV) is the archetype of the *Cardiovirus A* group within the family *Picornaviridae.* It has a 7.8 kb positive-sense, single-stranded, linear RNA genome comprising a single long open reading frame (ORF; ~2200 amino acids) flanked by an extended 5′ untranslated region (UTR) containing an internal ribosome entry site (IRES), and a shorter 3′ UTR with a poly(A) tail. Upon infection, the genome is translated directly to yield a polyprotein (L-1ABCD-2ABC-3ABCD) that is proteolytically processed into approximately 12

individual gene products by the viral 3C protease. In addition to IRES utilisation[1], the discovery of Stop-Go peptide release[2,3] and PRF[4,5] during genome translation has established EMCV as a model system for studying ribosome-related gene expression mechanisms. In EMCV, PRF occurs 11–12 codons into the start of the 2B gene, with up to 70% of ribosomes changing frame and producing the 2B* *trans*-frame product.

PRF is a translational control strategy employed by many RNA viruses, where it ensures the production of proteins in optimal ratios for efficient virus assembly and enables viruses to expand their coding capacity through the utilisation of overlapping ORFs (reviewed in[6-8]). In canonical PRF, elongating ribosomes pause over a heptanucleotide "slippery sequence" of the form X_XXY_YYZ when they encounter a "stimulatory element" 5–9 nucleotides downstream in the mRNA. During this time, a –1 frameshift may occur if codon-anticodon re-pairing takes place over the X_XXY_YYZ sequence: wobble positions allow the tRNA in the P-site tRNA to slip from XXY to XXX, and the tRNA in the A-site to slip from YYZ to YYY. Frameshifting may occur during a late stage of the EF-G/eEF2 catalysed translocation step, with the stimulatory element causing paused ribosomes to become trapped in a chimeric rotated or hyper-rotated state that is relieved by either the spontaneous unfolding of the blockade or a –1 slip on the mRNA[9-12]. A diverse array of stem-loops and pseudoknots are known to induce frameshifting, and the stability and unfolding kinetics of these stimulatory elements were initially thought to be the primary determinants of PRF efficiency[13,14]. However, more recently, the conformational plasticity of the elongation blockade has been revealed to play an important role[15-17]. Cardioviruses present a highly unusual variation to conventional viral PRF: the virally-encoded 2A protein is required as an essential *trans*-activator[5], and the stimulatory element is thought to comprise an RNA-protein complex formed between 2A and a stem-loop in the viral RNA[4]. This unique mechanism allows for temporal control of gene expression as the efficiency of –1 frameshifting is linked to 2A concentration, which increases with time throughout the infection cycle[4].

EMCV 2A is a small, basic protein (~17 kDa; 143 amino acids; pI ~9.1) generated by 3C-mediated proteolytic cleavage at the N-terminus[18] and Stop-Go peptide release at a C-terminal 18-amino acid consensus sequence[3]. Despite the identical name, the cardiovirus 2A has no homology to any other picornavirus "2A" protein[19], nor any other protein of known structure. Surprisingly, although cardiovirus replication and assembly is entirely cytoplasmic, 2A localises to nucleoli from early time-points post-infection[20,21]. As well as its role in stimulating PRF[4], 2A binds to 40S ribosomal subunits[22], inhibits apoptosis[23] and contributes to host cell shut-off by inhibiting cap-dependent translation, despite EMCV 2A differing from other picornavirus 2A proteins by having no protease activity against eIFs[24]. A previous mutational analysis[25] identified a putative nuclear localisation sequence (NLS) of the form $[G/P](K/R_3)x_{1-4}[G/P]$, similar to those found in yeast ribosomal proteins. This study also identified a C-terminal YxxxxLΦ motif, proposed to bind to and sequester eIF4E in a manner analogous to eIF4E binding protein 1 (4E-BP1), thereby inhibiting initiation of cap-dependent translation by interfering with eIF4F assembly[26]. Despite these insights, the absence of structural data has precluded a more definitive molecular characterisation of this multifunctional protein, and the mechanism by which it recognises RNA elements remains obscure.

Our previous RNA structure mapping experiments suggested that the stimulatory element in the EMCV genome adopts a stem-loop conformation, and we have demonstrated that a

conserved CCC motif in the putative loop region is essential for both 2A binding and PRF[4]. However, the nature of these tertiary interactions and the conformational dynamics of this frameshifting RNA element, including changes associated with 2A binding, are not well understood. For studying the thermodynamics and stability of these RNAs, the use of optical tweezers to conduct single-molecule force spectroscopy measurements can provide information beyond the resolution of conventional ensemble techniques, which are necessarily limited by molecular averaging[27]. In recent years, such approaches have yielded insights into various nucleic acid structures[28-31] and dynamic cellular processes[32,33] as well as mechanisms of PRF[16,34-36].

Here we present the crystal structure of EMCV 2A revealing a novel RNA-binding fold that we term a "beta-shell". Using a combination of biochemical and biophysical techniques, we show that 2A binds directly to the frameshift-stimulatory element in the viral RNA with nanomolar affinity and 1:1 stoichiometry, and we define the minimal RNA element required for binding. Furthermore, through site-directed mutagenesis and the use of single-molecule optical tweezers, we study the dynamics of this RNA element, both alone and in the presence of 2A. By observing short-lived intermediate states in real-time, we demonstrate that the EMCV stimulatory element exists in at least two conformations and 2A binding stabilises one of these, a putative RNA pseudoknot, increasing the force required to unwind it. Finally, we report a direct interaction of 2A with both mammalian and bacterial ribosomes. High-resolution cryo-electron microscopy (cryo-EM) characterisation of 2A in complex with initiated 70S ribosomes reveals a multivalent binding mechanism and defines the molecular basis for RNA recognition by the 2A protein. It also reveals a likely mechanism of 2A-associated translational modulation, by competing for ribosome binding with initiation factors and elongation factors. Together, our work provides a new structural framework for understanding protein-mediated frameshifting and 2A-mediated regulation of gene expression.

Results

Structure of EMCV 2A reveals a new RNA-binding fold

Following recombinant expression in *E. coli*, purified 2A was poorly soluble and prone to aggregation. Buffer screening by differential scanning fluorimetry[37] indicated that the thermal stability of the protein was enhanced by salt (data not shown), with high-salt buffers (~1M NaCl) greatly improving solubility. Size-exclusion chromatography coupled to multi-angle light scattering (SEC-MALS) revealed a predominantly monodisperse, monomeric sample (**Figure 1A and B**; observed mass 18032.8 Da vs 17930.34 Da calculated from the 2A sequence), with a small proportion of 2A forming dimers (observed mass 40836.0 Da). We crystallised the protein and, in the absence of a suitable molecular replacement search model, determined the structure by multiple-wavelength anomalous dispersion analysis of a selenomethionyl derivative. The asymmetric unit (ASU) of the $P6_222$ cell contains four copies of 2A related by non-crystallographic symmetry (NCS) and the structure was refined to 2.6 Å resolution (**Table 1**). Unexpectedly, the four molecules are arranged as a pair of covalent 'dimers' with an intermolecular disulfide bond formed between surface-exposed cysteine residues (C111). This arrangement is likely an artefact of crystallisation, which took >30 days, possibly due to the gradual oxidation of C111 promoting formation of the crystalline lattice. The N-terminal 10–12 residues are disordered in all chains except B, in which they make a long-range crystal contact with a symmetry-related molecule. Similarly, C-terminal residues beyond 137 are absent or poorly ordered in all chains.

2A adopts a compact, globular fold of the form $\beta_3\alpha\beta_3\alpha\beta$ (**Figure 1C**). Given the absence of structural homology to any other protein, we term this new fold a "beta shell". The most striking feature of this fold is a seven-stranded anti-parallel beta sheet that is highly curved (**Figure 1D**). The concave face of the beta sheet is supported by tight packing against the two alpha helices: together, this comprises the hydrophobic core of the fold. In contrast, the solvent-exposed convex face and surrounding loops are enriched with arginine, lysine and histidine residues, conferring a strong positive electrostatic surface potential at physiological pH. This is consistent with an RNA-binding mechanism in which the negatively charged ribose phosphate backbone is recognised by electrostatic interactions[38].

Superposition of the four NCS-related chains and an analysis of the atomic displacement factors reveals regions of flexibility within the 2A protein (**Figure 1E and F**). In addition to the N- and C- termini, the β2-loop-β3 region (residues 28–37) exists in multiple conformations that deviate by up to 5.8 Å in the position of the C_α backbone. Similarly, the arginine-rich loop between β5 and β6 ("arginine loop", residues 93–100) is highly flexible, with backbone deviations of up to 4.5 Å. Interestingly, this region has multiple roles: it acts as the 2A NLS[25] and mutation of R95 and R97 to alanine inhibits PRF by preventing 2A binding to the stimulatory element in the mRNA[4]. In support of the latter observation, we observe that this loop binds sulfate ions (present at high concentration in the crystallisation buffer) in two out of the four molecules in the ASU. Sulfate binding sites often indicate regions of a protein that could interact with an RNA phosphodiester backbone, based on similar geometry and charge.

Several previous studies have described mutations, truncations or deletions in EMCV 2A that affect its activity. We can now better understand the structural consequences of these alterations[25,39,40]. Many of the truncation mutants would lack substantial portions of secondary structure and expose elements of the 2A protein hydrophobic core (**Figure S1A and B**). This would severely disrupt the folding of the protein and the results obtained with these mutants should be interpreted with caution. However, the loop truncation (2A$_{\Delta94-100}$) and point mutations made by Groppo *et al.*[25] (**Figure S1C and D**) would not be predicted to disrupt the fold of 2A and can be interpreted in light of the structure. Notably, in 2A, a C-terminal YxxxxLΦ motif predicted to bind eIF4E is within a beta strand, whereas the equivalent motif in 4E-BP1 is alpha-helical. As a result of the more extended backbone conformation in 2A, Y129 is distal to L134 and I135. It is also partially buried and anchored in place by surrounding hydrophobic residues, in contrast to the tyrosine residue in 4E-BP1 that protrudes and makes significant contacts with a pocket on the eIF4E surface. Overlay of our 2A structure with the structure of the eIF4E:4E-BP1 complex indicates that, without a significant conformational change, this motif is unlikely to represent the mechanism by which 2A recognises eIF4E (**Figure S1E**).

2A binds to a minimal 47 nt pseudoknot in the viral RNA

The RNA sequence that directs PRF in EMCV consists of a G_GUU_UUU shift site, a variant of the canonical X_XXY_YYZ PRF slippery sequence, and a stimulatory stem-loop element downstream (**Figure 2A**). The spacing between shift-site and stem-loop is 13 nt, significantly longer than that seen typically (5–9 nt) at sites of –1 PRF, and 2A protein has been proposed to bridge this gap through interaction with the stem-loop. We have previously demonstrated that three conserved cytosines in the loop are essential for 2A binding[4] (**Figure 2A**). To map the interaction between 2A and the stimulatory element in more detail, we prepared a series of synthetic RNAs with truncations in the shift site, loop, and 5′ and 3′ extensions on either

side of the stem (EMCV 1–6; **Figure 2B**). These were fluorescently labelled at the 5′ end, and their binding to 2A was analysed by electrophoretic mobility shift assay (EMSA; **Figure 2C**) and microscale thermophoresis (MST; **Figure 2D and Table 2**).

Binding of 2A to EMCV 1 is high affinity (K_D = 360 ± 34 nM). This construct lacks the shift site, which would be within the ribosome and unavailable for 2A binding in a frameshift-relevant scenario. Removal of the 3′ extension, as in EMCV 3 and EMCV 6, further increases the affinity (K_D values of 40 ± 2 and 70 ± 14 nM, respectively), perhaps by removing competing base-pairing interactions. There is no substantial difference between affinities of EMCV 3 and 6, which differ only by the presence of the shift site. Removal of the 5′ extension, as in EMCV 2 and EMCV 4, completely abolishes 2A binding, and truncation of the loop, including a putative second stem (EMCV 5) reduces binding to micromolar levels. An EMSA was also performed with an N- and C-terminally truncated version of 2A containing a C111S mutation ($2A_{9-136; C111S}$), to probe whether the short peptide extensions added to the 2A N- and C-terminus during expression cloning or the disulfide bond observed in the crystal structure contribute to RNA binding. As seen (**Figure S2A**), this 2A variant bound EMCV 6 RNA identically compared to the wild-type protein. Inclusion of an N-terminal Strep-II tag (SII-2A) also had no effect on RNA binding (**Figure S2A**). In EMSAs of EMCV RNAs that bind 2A we also observe a lower-mobility species at higher protein concentrations, indicative of higher-order complex formation. To investigate the stoichiometry of binding, we performed isothermal titration calorimetry (ITC) analysis of the interaction between 2A and EMCV 6 (**Figure S2B and C**). Although the K_D of this reported interaction was higher (246 ± 72 nM) than observed using MST, possibly due to the higher salt concentration used to prevent 2A aggregation during the ITC experiment, the number of sites (0.87) is in good agreement with a 1:1 interaction. The largest contribution to the overall ΔG of binding (–9.02 kcal/mol) is the ΔH term (–13.9 ± 0.81 kcal/mol), consistent with an interaction mechanism driven by hydrogen bond or electrostatic contact formation. Finally, to test whether the presence of the fluorophore on the RNA affected 2A binding, we instead fluorescently labelled 2A and performed the reciprocal MST experiments with unlabelled RNA (**Figure S2D and Table 2**). The observed K_D values are in good agreement between the two approaches.

To further validate these observations, we asked whether the small EMCV stem-loop RNAs could act as competitors to sequester 2A and reduce the efficiency of PRF in rabbit reticulocyte lysate (RRL) *in vitro* translation reactions programmed with an EMCV dual luciferase frameshift reporter mRNA (**Figure S2E**). Indeed, when unlabelled EMCV 1, 3 and 6 were added in excess, they were able to compete with the stimulatory element present in the reporter, thereby reducing the amount of the –1 frame product. In contrast, EMCV 2, 4 and 5 had no such effect, reinforcing the results of direct binding experiments.

The failure of 2A to bind to EMCV 2, 4 and 5 was unexpected as these RNAs retain the main stem and the conserved cytosine triplet in the putative loop region. A possible explanation is that the frameshift-relevant state may include an interaction between the loop and the 5′ extension, forming a different conformation that 2A selectively recognises. Inspection of the primary sequences flanking the stem of the EMCV frameshift region revealed a number of possible base-pairing interactions, between 5′ or 3′ extensions and the loop, generating potential pseudoknots, and between the extensions themselves, generating an additional stem separated from the main stem by an internal loop. Whilst previous RNA structure probing

data[4] are largely consistent with the basic stem-loop model, we investigated the possibility that the EMCV PRF site forms a more complex structure by mutagenesis of the 5′ extension and loop C-triplet. Individually, G7C and C37G mutations both reduce 2A-dependent PRF to near-background levels (**Figure S3A and B**). However, in combination, the G7C+C37G double mutation restores PRF to wild-type levels, and EMSA experiments with these mutants confirm that this is due to inhibition and restoration of 2A binding (**Figure S3C**). Together, this demonstrates the likelihood of a base-pair between positions 7 and 37 that is necessary to form a conformation that 2A selectively recognises. Using this base pair as a restraint, RNA structure prediction[41,42] reveals a pseudoknot-like fold (**Figure S3D**).

Single-molecule measurements of stimulatory element unwinding reveal multiple states

We further explored the individual folding transitions within potential stem-loop and pseudoknot conformations by single-molecule force spectroscopy using optical tweezers (**Figure 3A**). We used the force-ramp method[43,44] to probe the force ranges of unfolding-refolding trajectories. Briefly, RNA molecules held between DNA handles were gradually stretched at a constant rate, and then the applied force was released, while recording the molecular end-to-end extension distances. This allows the molecule to transition between folded and unfolded states[43,44], and sudden changes in measured force-distance curves are indicative of transitions between RNA conformers. Alongside the wild-type EMCV RNA sequence, we also tested a mutant with a substitution in the cytosine triplet (CUC) known to inhibit 2A binding and PRF[4].

We initially monitored the unfolding and refolding of the wild-type (CCC) and mutant (CUC) RNAs in the absence of 2A protein. At increasing force, unfolding was observed in several steps representing different conformers (**Figure 4 and Table 3**). The most dominant state in the wild-type RNA was the predicted stem-loop, which was observed in 40% of the population. This state 1 (St1) unfolds in three steps (**Figure 4**), with the rips occurring at around 6, 12 and 25 pN. Upon release of the force, the molecule refolds with little perturbation. Overall, the unfolding and refolding behaviour of this population is consistent with an extended stem-loop model with internal loop (**Figure 3B**). The next major population, State 2 (St2; 20%), unfolds in two steps, with one small step (4–5 nm) occurring at low forces of around 8 pN followed by a full extension of 11 nm at 30 pN (**Figure 3C and 4B**). In contrast to the first population, this population (St2) has a different refolding behaviour with a large hysteresis, similar to previous observations on other known pseudoknot structures[14,15]. Refolding occurred at lower forces of ~15 pN and in some cases was not seen (**Figure 3C**, blue line). The third state (St3; 24%) represents a population in which unfolding occurred in a single low-force step with a small extension of around 6 nm, likely characteristic of a short stem loop (**Figure 4A and B**). Here, in contrast to the St1 and St2, no other unfolding steps were observed up to the maximum force applied (~40 pN). In a small fraction of the traces (St4), unfolding was observed in two low-force steps of 5 and 9–10 pN, which we predict may occur if the main stem does not fold properly. Finally, 8% of the traces showed no unfolding behaviour, even at high applied forces (St5).

Compared to wild-type, the main difference observed with the CUC RNA was the relative absence of the pseudoknot-like state (St2): only about 7% of the RNAs folded into this high force conformer. This finding is consistent with our biochemical data, suggesting that the cytosine triplet is involved in some long range, pseudoknot-like interactions with the 5'

extension (**Figure S3**). Instead, St1 and St3 states were observed in the majority of CUC traces (41% and 13%, respectively) and both displayed similar folding and unfolding transitions to equivalent states in CCC RNA. This is consistent with our expectation the predicted stem-loop would still be able to form in the CUC mutant (**Figure 3D**). St4 was observed in 10% of the population, occurring at similar forces (6 pN and 9 pN) and extensions (5 nm and 6 nm) to CCC RNA, and St5 was completely absent from the CUC population.

2A favours the formation of an alternative state with high resistance to mechanical unwinding

We next tested how 2A binding influences RNA stability and resistance to mechanical unwinding. For wild-type RNA, analysis of the frequency distribution of measured forces across all experiments reveals a global 2A-induced stabilisation, with increased numbers of observed high force (~25 pN) and very high force (>35 pN) unfolding events in the presence of 2A protein (**Figure 3E**). Within this population, we were able to identify the same five states from their unfolding and extension behaviour (**Figure 4A and B, Table 3**), yet the population densities showed significant differences compared to RNA-only experiments (**Figure 4C**). The proportions of predicted stem-loop and pseudoknot-like conformations (St1 and St2, respectively) were relatively unchanged by addition of 2A, but in these populations, we no longer observed a low-force step (8 pN and 6 nm) corresponding to the unfolding of short stems immediately 5′ to the main stem loop. Strikingly, the proportion of molecules in the low-force St3 and St4 states decreased in the presence of 2A (St3 24% to 9%; St4 10% to 1%), accompanied by a concomitant increase in the proportion of St5 (8% to 36%). St5 is highly resistant to unwinding, and we did not observe full extension even at forces of ~40 pN (**Figure 4B**).

Because St5 unfolds at forces beyond the maximum used in our experiments, we cannot determine whether the St5 conformers observed in the CCC and CCC+2A experiments are truly equivalent. It was also necessary to maintain a low concentration of 2A (~300 nM) to prevent aggregation and minimise non-specific interactions. Given our observed K_D values for 2A binding (**Figure 2D**), it is likely that a proportion of traces in CCC+2A experiments correspond to RNA-only events. In light of these caveats, several interpretations are possible. The simplest is that St5 is a distinct RNA conformation that exists in equilibrium with the others and that 2A binding stabilises this conformation, thus increasing its relative abundance. Alternatively, 2A may preferentially bind to semi-folded intermediates (St3 and St4), remodelling them into a highly stable state (St5) that differs from any conformation in the absence of 2A. Conversely, St5 could simply represent a 2A-bound and stabilised version of pseudoknot St2, with the same RNA conformation, but a higher unwinding force. In this scenario, St3 and St4 may be folding precursors to St2, and their disappearance as a result of 2A addition may be due to conversion to St2 as the equilibrium shifts towards St5 formation (i.e. St3-4 → St2 → St5). In this explanation, St2 would still be observed under non-saturating 2A concentrations.

Finally, we tested the effects of 2A on the CUC mutant RNA. Within this population we did not observe a large, global 2A-induced stabilisation (**Figure 3E**) and, unlike the wild-type RNA, the presence of 2A did not change the number of unfolding steps (**Table 3**). In addition, St5 was completely absent, in contrast to the CCC+2A experiments in which it is the major species. The observed proportions of St1 and St2 remained similar, and the low-force unfolding events in St3 showed a broader distribution, possibly due to non-specific interactions

between 2A and the handle regions. Together, our results suggest that 2A binding stabilises the stimulatory RNA element and increases its resistance to mechanical unwinding

2A interacts with eukaryotic and prokaryotic ribosomes

The high unwinding force of the 2A-bound St5 conformer likely reflects its role as the stimulatory element that induces a ribosomal pause at the PRF site[4,45]. However, in addition to its role as a component of the stimulatory element, 2A has been reported to bind to 40S subunits in EMCV-infected cells[22]. The direct interaction of 2A with ribosomes may be pertinent to its capacity to stimulate PRF: 2A may interact with translating ribosomes when they encounter the stimulatory 2A-RNA complex or (perhaps less likely) travel with elongating ribosomes to interact with the PRF signal. The 2A:40S interaction may also be relevant to the inhibition of host cell translation.

To determine if the interaction of 2A with the 40S subunit can be reproduced *ex vivo*, we purified ribosomal subunits from native RRL and analysed 2A-subunit interactions by MST (**Figures 5A and B**). Consistent with previous data, we were unable to detect an interaction with 60S, but 2A forms a tight complex with 40S (K_D = 10 ± 2 nM). To gain insight into this interaction, we prepared 2A-40S complexes for cryo-EM studies. Analysis by size-exclusion chromatography revealed that 2A co-eluted with the 40S peak (**Figure S4A and B**) but, despite extensive optimisation, subsequent cryo-EM imaging did not reveal interpretable density for 2A. As an alternative, we tested direct binding of 2A to purified prokaryotic 30S subunits by MST. 2A binds with very high affinity (K_D = 4 ± 1 nM; **Figure 5C**). We also examined binding of 2A to intact 70S ribosomes and to reconstituted, mRNA-bound 70S ribosomes at the initiation stage (70S IC; initiator tRNA[Met] in the P-site and an empty A-site). We were able to detect high affinity interactions with both uninitiated and initiated 70S ribosomes (**Figures 5D and E**). It is well established that prokaryotic translation systems are generally responsive to eukaryotic PRF signals[10,28,29,46,47] but this has not been tested for sites of protein-dependent PRF. To address this, we measured the efficiency of the EMCV signal in a reconstituted prokaryotic translation system (data not shown) and in *E. coli* S30 extracts using frameshift reporter mRNAs (**Figure S4C**). In each case, 2A-dependent PRF was observed, with ~ 15% of ribosomes changing frame. Shortening the length of the spacer to one more optimal for prokaryotic ribosomes (from 13 to 12 nt) led to a further two-fold increase in PRF. These efficiencies are comparable to those measured in eukaryotic *in vitro* translation systems (20%)[4] and high concentrations of 2A had an inhibitory effect on translation (**Figure S4D**), similar to that seen in eukaryotic systems.

Cryo-EM characterisation of a 2A-ribosome complex reveals the structural basis for RNA recognition and inhibition of translation

Given the high-affinity interaction, and having validated the use of prokaryotic ribosomes as a model system to study protein-dependent PRF, we prepared complexes between 2A and the initiated 70S ribosomes and imaged them by cryo-EM (**Figure 6A; Table 4**). After processing (**Figure S5A**), the final 3D reconstruction produced a density map of 2.7 Å resolution (**Figure S5B – D**) and revealed three copies of 2A bound directly to 16S rRNA of the 30S subunit in a tripartite cluster (**Figure 6B and C**). After docking the crystal structure (**Figure 1D**), the local resolution for 2A was high enough to allow sidechain modelling and refinement (**Figure 6D**). Alignment of the three RNA-bound conformations to the two main *apo-* conformations

observed in different NCS-related chains of the crystal structure shows that the 2A$_{apo}$1 (chain A-like) backbone conformation is more similar to the rRNA-bound state (**Figure S6A**).

All three copies of 2A bind directly to the ribose phosphate backbone via numerous polar and electrostatic contacts. 2A1 (orange) forms seven hydrogen bonds to rRNA helices 3, 4, 5,15 and 17, burying a surface of ~ 495 Å^2 (**Figure 6E**); 2A2 (red) makes 10 hydrogen bonds with helices 4, 15 and 17, burying a surface of ~ 606 Å^2 (**Figure 6F**) and 2A3 (yellow) forms 12 hydrogen bonds the backbone of helices 16 and 17, burying a surface of ~ 532 Å^2 (**Figure 6G**). In all three copies of 2A, the same RNA-binding surface is involved, comprising variations of residues R46, K48, K73, K50, K94, R95 and R97. Interestingly, the RNA binding targets differ between the 2A binding sites. The protein does not associate with regular helices; all of the targets contain regions of helical distortion or comprise helical junctions. The most important protein residues involved in RNA binding are R95, R97 and R100, present in the flexible "arginine loop" (**Figure 1F**). This loop adopts a different conformation in all three copies of 2A, allowing the arginine residues to bind to a wide variety of different RNA structures, not only via hydrogen bonding, but also via hydrophobic stacking interactions between exposed bases (G38) and arginine side chain guanidinium groups (**Figure 6H-J**). Whilst base-specific contacts are rare, 2A2 interacts with U485 which is normally flipped out of helix 17 (**Figure S6B**). Comparison of side-chain conformation at the RNA-binding surface in all three 2A molecules (**Figure 6K**) reveals a high-degree of conformational plasticity, explaining how this protein can recognise a diverse set of RNA molecules. There are also intermolecular contacts between 2A protomers. These interactions are consistent with our observations of multimers in both *apo-* and RNA-bound states by SEC-MALS (**Figure 1B**) and EMSA (**Figure 2C**), and the tendency for 2A to self-associate at physiological salt concentrations. 2A1 (orange) makes hydrophobic contacts with both other molecules, burying surfaces of ~ 423 Å^2 (2A2, red) and ~ 609 Å^2 (2A3, yellow). There are no direct interactions between 2A2 and 2A3, and none of the observed protein-protein interfaces resemble those seen in crystal contacts. The intermolecular disulfide bond present in the crystal lattice is also absent (**Figure 1E**).

The ribosome is in an unrotated state that would normally be elongation competent, with fMet-tRNA$_i$ base-paired to the initiator codon in the P-site and mRNA available for amino-acyl tRNA delivery to the A-site. There are no 2A-induced rearrangements at the decoding centre (**Figure S6C and D**) However, the presence of 2A on the 30S subunit occludes the binding site for translational GTPases. 2A1 occupies a position that would severely clash with domain II of EF-G in both compact and extended pre- and post-translocation states[48,49] (**Figure 6L**). It also makes direct hydrophobic contacts with the face of S12 that would normally interact with domain III of EF-G. This 2A interaction surface on S12 is directly adjacent to the binding site for antibiotic dityromycin, which inhibits translocation by steric incompatibility with the elongated form of EF-G[50] (**Figure S6E**). 2A1 would also clash significantly with domain II of EF-Tu during delivery of aminoacyl tRNAs to the A-site [51,52] (**Figure 6M**). In a similar way, 2A2 would be detrimental to both EF-G and EF-Tu binding (**Figure 6L and M**). We predict that this would have severe consequences for elongation. Indeed, at high levels, 2A is inhibitory to *in vitro* translation in both mammalian[4] and prokaryotic systems (**Figure S4D**). Binding at this site would also be inhibitory to initiation as it will compete for binding of IF2 during delivery of fMet-tRNA$_i$ to the P-site during pre-initiation complex assembly[53]. Conversely, 2A3 occupies a site that would not clash sterically with initiation or elongation factors. Given its role in PRF, we predicted that 2A may bind proximal to the mRNA entry channel close to 30S proteins associated with mRNA unwinding activity and decoding fidelity (S3, S4 and S5) but, despite

extensive focussed classification, no binding at this site was observed. However, a fourth copy of 2A (2A4) was identified to bind helix 33 of the 16S rRNA 'beak' in the 30S head (**Figure S5C-E**). Whilst the crystal structure could be unambiguously docked, the local resolution was insufficient for further modelling (**Figure S6E and F**). Nevertheless, 2A4 uses a similar binding mode to recognise the distorted helical backbone.

Discussion

Cardiovirus 2A is unique amongst picornaviral 2A proteins and a lack of homology to any known protein had precluded detailed functional inferences. Here we show that 2A adopts a novel RNA-binding fold, allowing specific recognition and stabilisation of the PRF stimulatory element in the viral RNA and direct binding to host ribosomes. The necessity for a functional Stop-Go motif at the 2A C-terminus has made a number of historical experiments difficult to interpret, as phenotypes may originate from impaired viral polyprotein processing rather than loss of specific 2A function[3,25]. Our structure therefore provides a framework to help rationalise several decades of preceding biochemical and virological observations.

Unusually for a multi-functional protein, it appears that many functions of 2A can be assigned to a single positively charged surface loop (residues 93–100). Despite the low pairwise sequence identity of 2A proteins amongst Cardioviruses (e.g. Theiler's murine encephalomyelitis virus [TMEV], Saffold virus, Rat theliovirus), R95 and R97 are completely conserved. This region was originally described as an NLS as mutation of these residues, or truncation of the whole loop, abolished 2A nuclear localisation[25]. Subsequently, we demonstrated that these residues are essential for PRF activity in both EMCV and TMEV, and that their mutation to alanine prevents 2A binding to the stimulatory element[4,45]. Here we reveal how R95 and R97 mediate direct 2A binding to the small ribosomal subunit (**Figure 6H – J**) and are therefore likely to play a critical role in conferring 2A-associated translational activities. The observed conformational heterogeneity of this loop (**Figures 1F and 6K**) indicates that mobility and flexibility are key to its myriad functions, particularly RNA binding.

Our cryo-EM structure unexpectedly revealed four distinct 2A:RNA interfaces (**Figure 6E – J and Figure S6D – E**), providing clues as to how RNA-binding specificity is achieved. RNA recognition is driven almost exclusively by electrostatic interactions between arginine or lysine side chains and the ribose phosphate backbone oxygen atoms; very few base-specific contacts are observed. Inspection of nearby RNA chains after superposition of the three well-resolved 2A molecules failed to reveal a common preferred backbone conformation, consistent with 2A being able to flexibly bind non-regular structured RNA including features such as kinks, distortions and junctions between multiple helices. Importantly, the 70S ribosome contains many examples of A-form helices with regular geometry but 2A does not bind to any of these sites. This is consistent with our experiments to define the minimal PRF stimulatory element in the viral mRNA (**Figure 2C and D**). Here, 2A is unable to bind EMCV 2, 4 and 5 RNAs, even though these constructs are predicted to form stable, undistorted stem-loops. Based on our biochemical data (**Figure S3**), there is a strong likelihood that, in the 2A-bound state, the conformation of the EMCV RNA that stimulates PRF involves additional base-pairs between C-residues in the loop and a GG pair in the 5′ extension. This pseudoknot-like conformation may either pre-exist in equilibrium with other states, or it may be directly induced by 2A binding (**Figure 4D**).

Our single-molecule data indicates that the conformational landscape of the EMCV PRF site is more complex than originally anticipated (**Figure 4A**). Besides the predicted stem-loop (St1) and pseudoknot conformations (St2), we also observed at least two other states with low-force unfolding steps. These are likely transition intermediates—partially folded or misfolded conformations of the predicted stem-loop or pseudoknot structures. On wild-type RNA, addition of 2A reduces the prevalence of these low energy states to background levels, and we see a major increase in the highly stable state 5, which may represent the 2A-bound pseudoknot-like conformation. This is accompanied by a global increase in unfolding forces across the entire population (**Figure 3E**). Conversely, on CUC mutant RNA, pseudoknot-like states (St2 and St5) do not form and the presence of 2A only induces a slight change in unfolding forces, which may result from non-productive interactions[54]. Moreover, we observe no 2A-induced differences in the distribution of unfolding pathways. This supports the idea that the failure of the CUC mutant to stimulate PRF is due to its inability to adopt pseudoknot-like conformations that would normally be selectively recognised or stabilised by the 2A.

Although the 2A protein favours the stabilisation of a distinct conformer in the wild-type RNA, in our model system this state co-exists with other predicted stem-loop-like and pseudoknot-like conformations. Comparison of measured force trajectories reveals that the 2A-bound state exhibits the greatest resistance to unwinding (St5, >35 pN; **Table 3**) and therefore may cause the longest ribosomal pause, providing an extended time window for frameshifting to occur. However, given that the maximum force the ribosome can generate during translocation on an mRNA is around 13–21 pN[55], a sufficient pause is also likely to be generated by other states (St1, ~25 pN; St2 ~29 pN; **Table 3**). It was originally thought that the higher energetic stability and thus slow unfolding kinetics are important for induction of PRF[14,29], however more recent studies report that rather than a static stability of the structure, a dynamic interplay between more conformations is crucial for efficient frameshifting[15-17]. Thus, the observed conformational heterogeneity at the EMCV PRF site may reflect a similar requirement for stimulatory element plasticity in protein-mediated PRF.

Our current mechanistic understanding of PRF is largely informed by ensemble kinetic and single-molecule FRET studies of prokaryotic ribosomes[9-11,56-58]. Frameshifting occurs late during the EF-G catalysed translocation step, in which the stimulatory element traps ribosomes in a rotated or hyper-rotated state, accompanied by multiple abortive EF-G binding attempts and rounds of GTP hydrolysis. A recent crystal structure showed that, in the absence of EF-G, tRNAs can spontaneously adopt a hybrid chimeric state with resultant loss of reading frame[59]. One model suggests that, in this state, an equilibrium is established between the 0 and −1 frame, which converges to 50% for long pause durations[11]. Based on our structure, it is tempting to speculate that competition between EF-G and 2A binding may have a role in further prolonging the pause, thereby contributing to the high PRF efficiencies that we observe in 2A-dependent systems[4,45]. However, the same residues in the 2A arginine loop are involved in binding both to the PRF stimulatory element in the viral RNA[4,45] and to ribosomal subunits (**Figure 6H – J**). Therefore, for any given molecule of 2A, these events are likely to be mutually exclusive. This implies that the ribosome-bound form of 2A that we observe could be a secondary 'enhancer' of PRF efficiency, acting synergistically with the main stimulatory element. It could also be relevant to the resolution of the elongation blockade: by providing an alternate 2A-binding surface that competes with the viral RNA, the ribosome may help to induce 2A dissociation from the stimulatory element during a pause at the PRF site.

Alternatively, it may not be directly relevant to frameshifting *per se*, instead representing a way of interfering with host cell translation as 2A accumulates later in infection. We cannot formally rule out the possibility that 2A3 (which does not occlude elongation factor binding, **Figure 6G**) may travel with the elongating ribosome and be 'unloaded' onto the PRF stimulatory element as the ribosome approaches, or may be involved in causing the ribosome to stall via protein-protein interactions between a ribosome-associated 2A and a stimulatory-element associated 2A. Indeed, the observation of a direct interaction between the ribosome and a PRF stimulatory element is not unprecedented, with a recent study revealing how the HIV-1 stem loop induces a stall by binding to the 70S A-site and preventing tRNA delivery[58]. Future kinetic studies and cryo-EM imaging of ribosomes advanced codon-by-codon along the mRNA may resolve this ambiguity.

Despite our structural insights, the precise mechanism by which 2A inhibits cap-dependent initiation remains enigmatic. In normal translation, a YxxxxLΦ motif in eIF4G mediates binding to eIF4E, thereby forming eIF4F and promoting initiation. 4E-BPs also contain a YxxxxLΦ motif, competing for eIF4E binding and acting as negative regulators[60]. A previous study proposed that a C-terminal YxxxxLΦ motif in 2A directly binds and sequesters eIF4E in a functionally analogous way to 4E-BP1[25], however our crystal structure suggests that this is unlikely to be the case without a drastic conformational rearrangement (**Figure S1E**). It is unclear how relevant the 2A-eIF4E interaction is to host cell shut-off, as viruses harbouring mutations in the putative YxxxxLΦ motif were still able to inhibit cap-dependent translation of host mRNAs despite losing the ability to bind eIF4E[25]. An alternative explanation is that binding of 2A to 40S inhibits translation initiation. Although our cryo-EM structure suggests that 2A may block binding of IF2 to the 30S subunit, prokaryotic initiation is significantly different to cap-dependent eukaryotic initiation. Even if a similar mechanism did occur to prevent eIF2 binding in eukaryotes, this would also inhibit viral type II IRES-mediated initiation which requires all initiation factors except eIF1, eIF1A and intact eIF4F[61,62]. In future it will be informative to further dissect the involvement of 2A in both IRES-dependent and cap-dependent translational initiation.

In conclusion, this work defines the structural and molecular basis for the temporally regulated 'switch' behind the reprogramming of viral gene expression in EMCV infection (**Figure 7**). At the heart of this is the 2A protein: a novel RNA-binding fold with the remarkable ability to discriminate between stem-loop and pseudoknot conformers of the PRF stimulatory element. We also reveal how 2A interferes with host translation by specifically recognising distinct conformations within the ribosomal RNA. Together, this illustrates how the conformational plasticity of one RNA-binding surface can contribute to multiple functions through finely tuned relative affinities for different cellular targets.

Materials and Methods

Protein expression and purification

EMCV 2A cDNA was amplified by PCR from previously described plasmid pGEX6P1[4] (Forward: 5′ AATTCATATGAGCCCCAACCCTTTGGATGTC 3′, Reverse: 5′ AATTGGATCCCCCGGGATTGGTCTCGACATC 3′) and cloned into pOPTnH[63] using NdeI and BamHI sites, thereby introducing a C-terminal GlySerLysHis$_6$ tag. Protein was expressed in *E. coli* BL21(DE3) pLysS cells, grown with shaking (210 rpm, 37°C) in 2xTY broth

supplemented with 100 μg/mL ampicillin to an A_{600} of ~ 1 prior to induction of expression (210 rpm, 21°C, 16h) with 1.0 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). For selenomethionyl derivatisation ($2A_{SeMet}$), protein was expressed in *E. coli* B834 cells, grown shaking (210 rpm, 37°C) in SeMet base media (Molecular Dimensions) supplemented with nutrient mix, 40 μg/mL L-selenomethionine and 100 μg/mL ampicillin. Expression was induced as above.

Cells were harvested by centrifugation (4,000 × g, 4°C, 20 min), washed once in ice-cold PBS and stored at -20°C. Pellets from four litres of culture were resuspended in cold lysis buffer (50 mM Tris-HCl pH 8.0, 500 mM NaCl, 30 mM imidazole, supplemented with 50 μg/mL DNase I and EDTA-free protease inhibitors) and lysed by passage through a cell disruptor at 24 kPSI (Constant Systems). Lysate was cleared by centrifugation (39,000 × g, 40 min, 4°C) prior to incubation (1 h, 4°C) with 4.0 mL of Ni-NTA agarose (Qiagen) pre-equilibrated in the same buffer. Beads were washed in batch four times with 200 mL buffer (as above, but without DNase or protease inhibitors) by centrifugation (600 × g, 10 min, 4°C) and re-suspension. Washed beads were pooled to a gravity column prior to elution over 10 column volumes (CV) with 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 300 mM imidazole. Fractions containing 2A were pooled and dialysed (3K molecular weight cut-off (MWCO), 4°C, 16 h) against 1 L buffer A (50 mM Tris-HCl pH 8.0, 400 mM NaCl, 5.0 mM DTT) before heparin-affinity chromatography to remove contaminating nucleic acids. Samples were loaded on a 10 mL HiTrap Heparin column (GE Healthcare) at 2.0 mL/min, washed with two CV of buffer A and eluted with a 40% → 100% gradient of buffer B (50 mM Tris-HCl pH 8.0, 1.0 M NaCl, 5.0 mM DTT) over 10 CV. Fractions containing 2A were pooled and concentrated using an Amicon® Ultra centrifugal filter unit (10K MWCO, 4,000 × g). Size exclusion chromatography was performed using a Superdex 75 16/600 column pre-equilibrated in 10 mM HEPES pH 7.9, 1.0 M NaCl, 5.0 mM DTT. Purity was judged by 4-20% gradient SDS-PAGE, and protein identity verified by mass spectrometry. Purified protein was used immediately or was concentrated as above (~ 7.0 mg/mL, 390 μM), snap-frozen in liquid nitrogen and stored at -80°C. Variants of 2A, including $2A_{9-136;C111S}$ and $2A_{SeMet}$ were purified identically to the wild-type protein.

Size-exclusion chromatography coupled to multi-angle light scattering (SEC-MALS)

Per experiment, 100 μL of protein was injected onto a Superdex 75 increase 10/300 GL column (GE Healthcare) pre-equilibrated with 20 mM Tris-HCl, 1.0 M NaCl (0.4 mL/min flow, 25°C). Experiments were performed with 5.2 mg/mL 2A (corresponding to a molar concentrations of 290 μM). The static light scattering, differential refractive index, and the UV absorbance at 280 nm were measured in-line by DAWN 8+ (Wyatt Technology), Optilab T-rEX (Wyatt Technology), and Agilent 1260 UV (Agilent Technologies) detectors. The corresponding molar mass from each elution peak was calculated using ASTRA 6 software (Wyatt Technology).

Protein crystallization

Purified EMCV 2A was concentrated to 5.9 mg/ml in 10 mM HEPES pH 7.9, 1.0 M NaCl, 2.0 mM DTT. Diffraction-quality native 2A crystals were grown at 21°C by sitting-drop vapor diffusion against an 80 μL reservoir of 0.625 M $(NH_4)_2SO_4$, 0.15 M tri-sodium citrate pH 5.7. Notably, crystal growth was only visible after 30 days. Drops were prepared by mixing 200 nL protein and 200 nL crystallization buffer. Selenomethionyl derivative 2A ($2A_{SeMet}$) was

concentrated to 5.7 mg/mL in 10 mM HEPES pH 7.9, 1.0 M NaCl, 2.0 mM DTT, and diffraction-quality 2A$_{SeMet}$ crystals were grown as above against an 80 µL reservoir of 0.675 M $(NH_4)_2SO_4$, 0.15 M tri-sodium citrate pH 5.7. Crystals were cryo-protected by the addition of 0.5 µL crystallization 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

Native datasets (**Table 1**) of 900 images were recorded at Diamond Light Source, beamline I03 (λ = 0.9796 Å) on a Pilatus 6M detector (Dectris), using 100% transmission, an oscillation range of 0.2° and an exposure time of 0.04 s per image. Data were collected at a temperature of 100 K. Data were processed with the XIA2[64] automated pipeline, using XDS[65] for indexing and integration, and AIMLESS[66] for scaling and merging. Resolution cut-off was decided by a $CC_{1/2}$ value ≥ 0.5 and an I/σ(I) ≥ 1.0 in the highest resolution shell[67]. For multiple-wavelength anomalous dispersion (MAD) phasing experiments, selenomethionyl derivative datasets were recorded at beamline I03 (peak λ = 0.9796 Å, 12656.0 eV; hrem λ = 0.9763, 12699.4 eV; inflexion λ = 0.9797 Å, 12655.0 eV). Data were processed as above using XIA2, XDS and AIMLESS. The structure was solved by three-wavelength anomalous dispersion analysis of the selenium derivative (space group $P6_222$) performed using the autoSHARP pipeline[68], implementing SHELXD[69] for substructure determination, SHARP for heavy-atom refinement and phasing, SOLOMON[70] for density modification and ARP/wARP[71] for automated model building. This was successful in placing 503/573 (87%) residues in the asymmetric unit, which comprised four copies of the protein related by non-crystallographic symmetry (NCS). This initial model was then used to solve the native dataset by molecular replacement with Phaser[72]. The model was completed manually by iterative cycles of model-building using COOT[73] and refinement with phenix.refine[74], using local NCS restraints. MolProbity[75] was used throughout the process to evaluate model geometry. For the electrostatic potential calculations, partial charges were first assigned using PDB2PQR[76], implementing PROPKA to estimate protein pKa values. Electrostatic surfaces were then calculated using APBS[77]. Prior to designation of the "beta shell" as a new fold, structure-based database searches for proteins with similar folds to EMCV 2A were performed using PDBeFOLD[78], DALI[79] and CATHEDRAL[80]. Buried surface areas were calculated using PDBePISA[81].

Electrophoretic Mobility Shift Assay (EMSA)

Synthetic RNA oligonucleotides (IDT) were dissolved in distilled water. RNAs were labelled at the 5′ end with A647-maleimide or Cy5-maleimide conjugates (GE Healthcare) using the 5′ EndTag kit (Vector Labs) as directed by the manufacturer. For each binding experiment, a series of reactions were prepared on ice, each containing 1.0 µL 500 nM RNA, 1.0 µL serially diluted protein at concentrations of 320, 160, 80, 40, 20, 10, 5.0, and 2.5 µM in 10 mM HEPES pH 7.9, 1.0 M NaCl, 5.0 µL 2 × buffer (20 mM Tris-HCl pH 7.4, 80 mM NaCl, 4.0 mM magnesium acetate 2.0 mM DTT, 10% v/v glycerol, 0.02% w/v bromophenol blue, 200 µg/mL porcine liver tRNA, 800 U /mL SUPERase-In [Invitrogen]) and 3.0 µL distilled water. This gave final binding reactions of 10 µL with 50 nM RNA, 1 × buffer, a salt concentration of ~ 140 mM and proteins at concentrations of 32, 16, 8.0, 4.0, 2.0, 1.0, 0.5 and 0.25 µM. Samples were incubated at 37°C for 20 min prior to analysis by native 10% acrylamide/TBE PAGE (25 min, 200 V constant). Gels were scanned with a Typhoon FLA-7000 (GE) using the 635 nm laser / R670 filter.

Isothermal Titration Calorimetry (ITC)

ITC experiments were performed at 25°C using an automated MicroCal PEAQ-ITC platform (Malvern Panalytical). Proteins and synthetic RNA oligonucleotides (IDT) were dialysed extensively (24 h, 4°C) into buffer (50 mM Tris-HCl pH 7.4, 400 mM NaCl) prior to experiments. RNA (52 µM) was titrated into protein (5 µM) with 1 x 0.4 µL injection followed by 12 × 3.0 µL injections. Control titrations of RNA into buffer, buffer into protein and buffer into buffer were also performed. Data were analysed using the MicroCal PEAQ-ITC analysis software (Malvern Panalytical) and fitted using a one-site binding model.

Microscale Thermophoresis (MST)

For RNA-binding experiments, synthetic EMCV RNA variants (IDT) were dissolved in distilled water and labelled with Dylight 650 maleimide conjugates (Thermo Scientific) as described above. RNA was diluted to 10 nM in MST buffer (50 mM Tris-HCl pH 7.8, 150 mM NaCl, 10 mM MgCl$_2$, 2 mM DTT supplemented with 0.05% Tween 20) and a series of 16 1:1 dilutions was prepared using the same buffer, producing ligand concentrations ranging from 0.00015 to 5 µM of EMCV 2A protein for RNA 2-6 and 0.0006 to 20 µM for RNA1. For the measurement, each ligand dilution was mixed with one volume of labelled RNA, which led to a final concentration of labelled RNA of 5.0 nM. The reaction was mixed by pipetting, incubated for 10 min followed by centrifugation at 10 000 × g for 10 min, the samples were loaded into Monolith NT.115 Premium Capillaries (NanoTemper Technologies). Measurements were performed using a Monolith NT.115Pico instrument (NanoTemper Technologies) at an ambient temperature of 25°C. Instrument parameters were adjusted to 5% LED power, medium MST power and MST on-time of 10 seconds. Data of two independently pipetted measurements were analysed for the Fnorm (MO.Affinity Analysis software, NanoTemper Technologies) and fraction bound and affinity constants were calculated using GraphPad Prism 8.0.2 software.

Conjugation of a fluorescent label to the surface-exposed cysteine residue (C111) observed in the 2A crystal structure (**Figure 1E**) provided a convenient way of studying binding to multiple unlabelled targets by MST, in such a way that the observed affinities would be directly comparable. EMCV 2A protein was labelled using the Protein Labelling Kit RED-Maleimide (NanoTemper Technologies) according to the manufacturer's instructions. In brief, 2A protein was diluted in a buffer containing 10 mM HEPES pH 7.9, 1.0 M NaCl and dye was mixed at a 1:3 molar ratio at room temperature for 30 min in the dark. Unreacted dye was removed on a spin gel filtration column equilibrated with 10 mM HEPES pH 7.9, 1.0 M NaCl. The labelled 2A protein was adjusted to 10 nM with MST buffer. Synthetic EMCV RNA variants were dissolved in the same buffer conditions, and a series of 16 1:1 dilutions was prepared using the same buffer, producing ligand concentrations ranging from 0.0008 to 26 µM for RNA 1 and 0.00003 to 1 µM for RNA 2-6. For the measurement, each ligand dilution was mixed with one volume of labelled protein 2A, which led to a final concentration of Protein 2A of 5.0 nM. Similar experiments were conducted for the ribosomes, with ligand concentrations 0.00002 to 0.4 µM for 40S and 60S, 0.00003 to 1 µM for 30S, 0.0008 to 1.375 µM for 70S and 0.000003 to 0.1 µM for 70S IC. The measurements were performed as described above.

Preparation of constructs for optical tweezer experiments

DNA encoding the frameshifting sequence of EMCV was inserted into plasmid pMZ_lambda_OT using PCR and subsequent Gibson assembly. This plasmid contains the ColE1 origin, ampicillin resistance, ribosome binding site and two 2 kbp handle regions derived from lambda phage DNA (5′ and 3′ handle). For the generation of the mutant plasmid, PCR and blunt-end ligation was used to mutate the CCC triplet in the EMCV stem-loop to CUC. Wild-type and mutant plasmids were subsequently used to generate construct suitable for optical tweezer measurements consisting of the EMCV frameshifting sequence flanked by the 2 kbp long handle regions. Three pairs of primers for PCR were designed allowing the amplification of the *in vitro* transcription template and 5′ and 3′ handles. Subsequently, PCR reactions generated 5′ and 3′ handles and a long template for *in vitro* transcription. The 3′ handle was labelled during PCR using a 5′ digoxigenin-labelled reverse primer. The 5′ handle was labelled with Biotin-16-dUTP at the 3′ end following PCR using T4 DNA polymerase. RNA was transcribed from templates for *in vitro* transcription using T7 RNA polymerase. RNA and both DNA handles (5′ and 3′) were annealed together in a mass ratio 1:1:1 (5 µg each) by incubation at 95 °C for 10 min, 62 °C for 1 hour, 52 °C for 1 hour and slow cooling to 4 °C in a buffer containing 80% formamide, 400 mM NaCl, 40 mM HEPES, pH 7.5, and 1 mM EDTA[82]. Following annealing, the samples were concentrated by ethanol precipitation, the pellets resuspended in 40 µl RNase-free water, split into 4 µl aliquots and stored at –20°C.

Optical tweezer data collection and analysis

Optical tweezer experiments were performed using a commercial dual-trap platform equipped with a microfluidics system (C-trap, Lumicks). An aliquot of the Optical tweezers (OT) construct was mixed with 1 µl of polystyrene beads coated with antibodies against digoxigenin (0.1% v/v suspension, Ø 1.76 µm, Lumicks), 5 µl of measurement buffer (20 mM HEPES, pH 7.6, 300 mM KCl, 5 mM MgCl$_2$, 5 mM DTT and 0.05% Tween) and 0.5 µl of RNase inhibitors. The mixture was incubated for 20 min at room temperature in a final volume of 10.5 µl, and subsequently diluted by addition of 0.5 ml measurement buffer. Separately, 0.8 µl of streptavidin-coated polystyrene beads (1% v/v suspension, Ø 2 µm, Lumicks) was supplemented with 1 ml of measuring buffer, the measuring cell was washed with the measuring buffer and suspensions of both streptavidin beads as well as the complex of OT construct with antidigoxigenin beads were introduced into the measuring cell.

Per experiment, an antidigoxigenin (AD) bead and a streptavidin (SA) bead were optically trapped and brought into close proximity to allow the formation of a tether in between. The beads were moved apart (unfolding) and back together (refolding) at constant speed (0.05 µm/s) to yield the force-distance (FD) curves. The stiffness was maintained at 0.3 and 0.24 pN/nm for trap 1 (AD bead) and trap 2 (SA bead), respectively. For experiments with 2A protein experiments, protein was diluted to 300 nM in measuring buffer and added to the buffer channel of the optical tweezer measuring cell. FD data was recorded at a rate of 78000 Hz and afterwards was filtered using Butterworth filter for down sampling by a factor of 20 (0.05 filtering frequency). Individual unfolding/refolding steps were detected by custom written algorithms using Matlab software. Custom written scripts were used to generate the histogram data. The FD curves were plotted using Prism 8 software. The RNAstructure software (version 6.2) was used for the prediction of the EMCV RNA element secondary structure[83].

Eukaryotic ribosomal subunit purification

40S and 60S subunits were purified from untreated rabbit reticulocyte lysate (Green Hectares) as previously described[44]. Briefly, ribosomes were pelleted by centrifugation (4°C, 270,000 × g, 4.5 h) and resuspended in 20 mM Tris-HCl pH 7.5, 4.0 mM $MgCl_2$, 50 mM KCl, 2.0 mM DTT. Following treatment with 1.0 mM puromycin and addition of KCl to 0.5 M, 40S and 60S subunits were separated by centrifugation (4°C, 87,000 × g, 16 h) through a sucrose density gradient (10 → 30% sucrose in 20 mM Tris-HCl pH 7.5, 2.0 mM DTT, 4.0 mM $MgCl_2$, 0.5 M KCl). After analysis by SDS-PAGE, uncontaminated fractions were pooled, and exchanged into 20 mM Tris-HCl pH 7.5, 100 mM KCl, 2.0 mM $MgCl_2$, 2.0 mM DTT, 250 mM sucrose using Amicon centrifugal concentrators (4°C,100K MWCO). Ribosome subunits were snap-frozen in liquid nitrogen and stored at -80°C until required.

Analytical SEC

Ribosome subunits were diluted into 20 mM Tris-HCl pH 7.5, 100 mM potassium acetate, 2.0 mM DTT, 1.5 mM $MgCl_2$, 0.25 mM spermidine to final concentrations of 0.5 – 1.5 µM. 2A protein was added in 10-fold molar excess and the mixture incubated on ice for 10 min prior to injection (83 µL, 0.04 mL/min, 4°C) onto a Superose 6 increase 3.2/300 column (GE) pre-equilibrated in 20 mM Tris-HCl pH 7.5, 100 mM potassium acetate, 2.0 mM DTT, 1.5 mM $MgCl_2$.

Western blot

Fractions from analytical size-exclusion chromatography were analysed by 4–20% gradient SDS-PAGE and transferred to a 0.2 µm nitrocellulose membrane. All subsequent steps were carried out at room temperature. Membranes were blocked (5% w/v milk, PBS, 1 h) before incubation (1 h) with primary antibodies in 5% w/v milk, PBS, 0.1% v/v Tween-20. Membranes were washed three times with PBS, 0.1% v/v Tween-20 prior to incubation (1 h) with IRDye fluorescent antibodies in 5% w/v milk, PBS, 0.1% v/v Tween-20. After three washes in PBS, 0.1% v/v Tween-20 and a final rinse in PBS, membranes were imaged using an Odyssey CLx Imaging System (LI-COR). Antibodies used were rabbit polyclonal anti-2A[4] (1/1000); mouse monoclonal anti-RPS6 (1/1000, clone A16009C, BioLegend); mouse monoclonal anti-RPL4 (1/1000, clone 4A3, Sigma); goat anti-rabbit IRDye 800 CW (1/10,000, LI-COR) and goat anti-mouse IRDye 680LT (1/10,000, LI-COR).

In vitro transcription

For *in vitro* frameshifting assays, we cloned a 105 nt DNA fragment containing the EMCV slippery sequence flanked by 6 nt upstream and 92 nt downstream into the dual luciferase plasmid pDluc at the XhoI and BglII sites[84]. This sequence was inserted between the Renilla and firefly luciferase genes such that firefly luciferase expression is dependent on −1 PRF. Wild-type or mutated frameshift reporter plasmids were linearized with FspI and capped run-off transcripts generated using T7 RNA polymerase as described[85]. Messenger RNAs were recovered by phenol/chloroform extraction (1:1 v/v), desalted by centrifugation through a NucAway Spin Column (Ambion) and concentrated by ethanol precipitation. The mRNA was resuspended in water, checked for integrity by agarose gel electrophoresis, and quantified by spectrophotometry.

Messenger RNAs for 70S IC preparation were produced from a 117 nt long DNA fragment containing the EMCV frameshift site flanked by the bacterial 5′ UTR with Shine-Dalgarno sequence and 18 nt downstream region of the putative structure.

5′GGGAAUUCAAAAAUUGUUAAGAAUUAAGGAGAUAUACAU<u>AUG</u>GA**GGUUUUU**AUCACUCAA
GGAGCGGCAGUGUCAUCAAUGGCUCAAACCCUACUGCCGAACGACUUGGCCAGATCT 3′
(slippery sequence in bold, initiator codon underlined)

This sequence was PCR amplified and *in vitro* transcribed using T7 RNA polymerase (produced in-house). Messenger RNAs were purified using the Qiagen RNeasy midiprep kit according to the manufacturer's protocols. The mRNAs were eluted in RNAse-free water, integrity and purity was checked by gel electrophoresis and quantified by spectrophotometry.

70S initiation complex preparation

Ribosomes, translation factors, and tRNAs were of *E. coli* origin. Total *E. coli* tRNA was from Roche, and oligonucleotides were from Microsynth. 70*S* ribosomes from MRE600, EF-Tu, EF-G, IF1, IF2 and IF3 were purified from *E. coli*[86]. fMet-tRNA[fMet] was prepared and aminoacylated according to published protocols[87,88]. Aminoacylated fMet-tRNA[fMet] was purified by reversed-phase HPLC on a Wide Pore C5 (10 µM particle size 10 mm x 25 cm) column (Sigma Aldrich). To prepare initiation complexes, 70S ribosomes (1 µM) were incubated with a three-fold excess of an EMCV model mRNA encoding for 5′…AUGGA**GGUUUUU**AUC…3′ (slippery sequence in bold) and a 1.5- fold excess each of IF1, IF2, IF3, and fMet-tRNA[fMet] in buffer A (50 mM Tris-HCl pH 7.5, 70 mM NH$_4$Cl, 30 mM KCl, 7 mM MgCl$_2$) supplemented with GTP (1 mM) for 30 min at 37°C. 70S initiation complexes were purified by centrifugation through a 1.1 M sucrose cushion in buffer A. Before grid preparation, initiation complexes were additionally purified on Sephacryl S-300 gel filtration microspin columns.

In vitro translation

Messenger RNAs were translated in nuclease-treated rabbit reticulocyte lysate (RRL) or wheat germ (WG) extracts (Promega). Typical reactions were composed of 90% v/v RRL, 20 µM amino acids (lacking methionine) and 0.2 MBq [^{35}S]-methionine and programmed with ~50 µg/ml template mRNA. Reactions were incubated for 1 h at 30°C. Samples were mixed with 10 volumes of 2× Laemmli's sample buffer, boiled for 3 min and resolved by SDS-PAGE. Dried gels were exposed to a Storage Phosphor Screen (PerkinElmer) and the screen scanned in a Typhoon FLA7000 using phosphor autoradiography mode. Bands were quantified using ImageQuant™TL software. The calculations of frameshifting efficiency (%FS) took into account the differential methionine content of the various products and %FS was calculated as % −1FS = 100 × (IFS/MetFS) / (IS/MetS + IFS/MetFS). In the formula, the number of methionines in the stop and frameshift products are denoted by MetS, MetFS respectively; while the densitometry values for the same products are denoted by IS and IFS respectively. All frameshift assays were carried out a minimum of three times.

Ribosomal frameshift assays in *E. coli* employed a coupled T7/S30 *in vitro* translation system (Promega). A ~450 bp fragment containing the EMCV PRF signal (or mutant derivative) was prepared by PCR from plasmid pDluc/EMCV[4] and cloned into the BamHI site of the T7-based, *E. coli* expression vector pET3xc[89]. T7/S30 reaction mixes were prepared according to the

manufacturer's instructions (50μl volumes), including 10 μCi ^{35}S methionine, supplemented with plasmid DNA (4 μg) and incubated at 37 ˚C for 90 mins. Reactions were precipitated by addition of an equal volume of acetone, dissolved in Laemmli's sample buffer and aliquots analysed by SDS-PAGE. PRF efficiencies were calculated as above.

Cryo-EM specimen preparation

Initiated 70S ribosomes in 50 mM Tris-HCl pH 7.5, 70 mM NH$_4$Cl, 30 mM KCl, 7 mM MgCl$_2$ were diluted tenfold into 20 mM HEPES pH 7.5, 100 mM potassium acetate, 1.5 mM MgCl$_2$, 2.0 mM DTT. 2A protein was dialysed (3K MWCO, 4°C, 16 h) into the same buffer. Crosslinking reactions of 50 μL comprising 75 nM ribosomes, 3.0 μM 2A and 2.0 mM bis(sulfosuccinimidyl)suberate (BS3) were performed on ice (30 min) immediately prior to grid preparation. Quantifoil R 2/2 400-mesh copper supports were coated with an additional ~ 60 Å layer of amorphous, evaporated carbon by flotation[90], and thoroughly dried before use. Grids were made hydrophilic by glow-discharge in air for 30 s. Three microliters of crosslinking reaction was applied to grids which were then blotted for 4.5 s and vitrified by plunging into liquid ethane using a Vitrobot MK IV (FEI) at 4°C, 100% relative humidity.

Cryo-EM data collection and processing

Micrographs were collected at the BiocEM facility (Department of Biochemistry, University of Cambridge) on a Titan Krios microscope (FEI) operating at 300 kV and equipped with a Falcon III detector (**Table 4**). At 75,000 × magnification, the calibrated pixel size was 1.07 Å / pixel. Per 0.6 s acquisition in integration mode, a total exposure of 54.4 e$^-$ / Å^2 was fractionated over 23 frames with applied defocus of –1.5, –1.8, –2.1, –2.4, –2.7 and –3.0 μm. EPU software was used for automated acquisition with five images per hole. After manual inspection, 5730 micrographs were used in subsequent image processing.

Movie frames were aligned and a dose-weighted average calculated with MotionCor2[91] The contrast transfer function (CTF) was estimated using CtfFind4[92]. All subsequent image-processing steps were carried out in RELION 3.1[93] (**Fig S5**) and all reported estimates of resolution are based on the gold standard Fourier shell correlation (FSC) at 0.143, and the calculated FSC is derived from comparisons between reconstructions from two independently refined half-sets. Reference-free autopicking of 820,475 particles was performed using the Laplacian-of-Gaussian function (200 - 250 Å diameter). Particles were initially downscaled threefold and extracted in a 150-pixel box. Two rounds of 2D classification (into 100 and 200 classes, respectively) were used to clean the dataset to 750,029 'good' particles. An initial model was generated from a PDB file of a 70S elongation-competent ribosome (PDB ID 6O9J) and low-pass filtered to 80 Å resolution. The initial 3D refinement (6.5 Å resolution) showed clear evidence for at least one copy of 2A adjacent to the factor binding site on the 30S subunit. At this stage, two rounds of focussed classification with signal subtraction were performed (6 classes) to separate particles based on additional density near i) the factor binding site and ii) the mRNA entry channel/helicase. The former was successful and 289,741 particles containing three copies of 2A were rescaled to full size and extracted in a 450-pixel box. Following initial 3D refinement, creation of a 15 Å low-pass filtered mask (five-pixel extension and five-pixel soft edge) and post-processing, a reconstruction of 2.93 Å was achieved. After per-particle CTF refinement and polishing, this was increased to 2.50 Å. With the increased angular accuracy provided by the fully rescaled data, focussed classification with signal

subtraction and local angular searches was performed again to separate particles based on 2A occupancy at the factor binding site. The final reconstruction (2.66 Å) from 120,749 particles revealed three copies of 2A bound with full occupancy. Calculation of a local resolution map revealed additional low-resolution density adjacent to the beak of the 30S head. Subsequent focussed classification with signal subtraction and refinement confirmed that this was a fourth copy of 2A bound, present in 73,059 particles.

Visualisation of structural data

All structural figures depicting crystallographic data (cartoon, stick and surface representations) were rendered in PyMOL (Schrödinger LLC). Structural figures of EM maps with docked components were rendered in ChimeraX[94].

Author contributions

C.H.H. and S.N. cloned expressed and purified proteins and performed all biochemical experiments. C.H.H. and S.C.G performed crystallography experiments. A.K. and N.C. performed MST experiments. L.P. and N.C. performed single-molecule experiments and analyses. C.H.H. prepared cryo-EM grids, and collected and processed cryo-EM data. C.H.H., S.N., N.C. and I.B. wrote the manuscript with contributions from all authors.
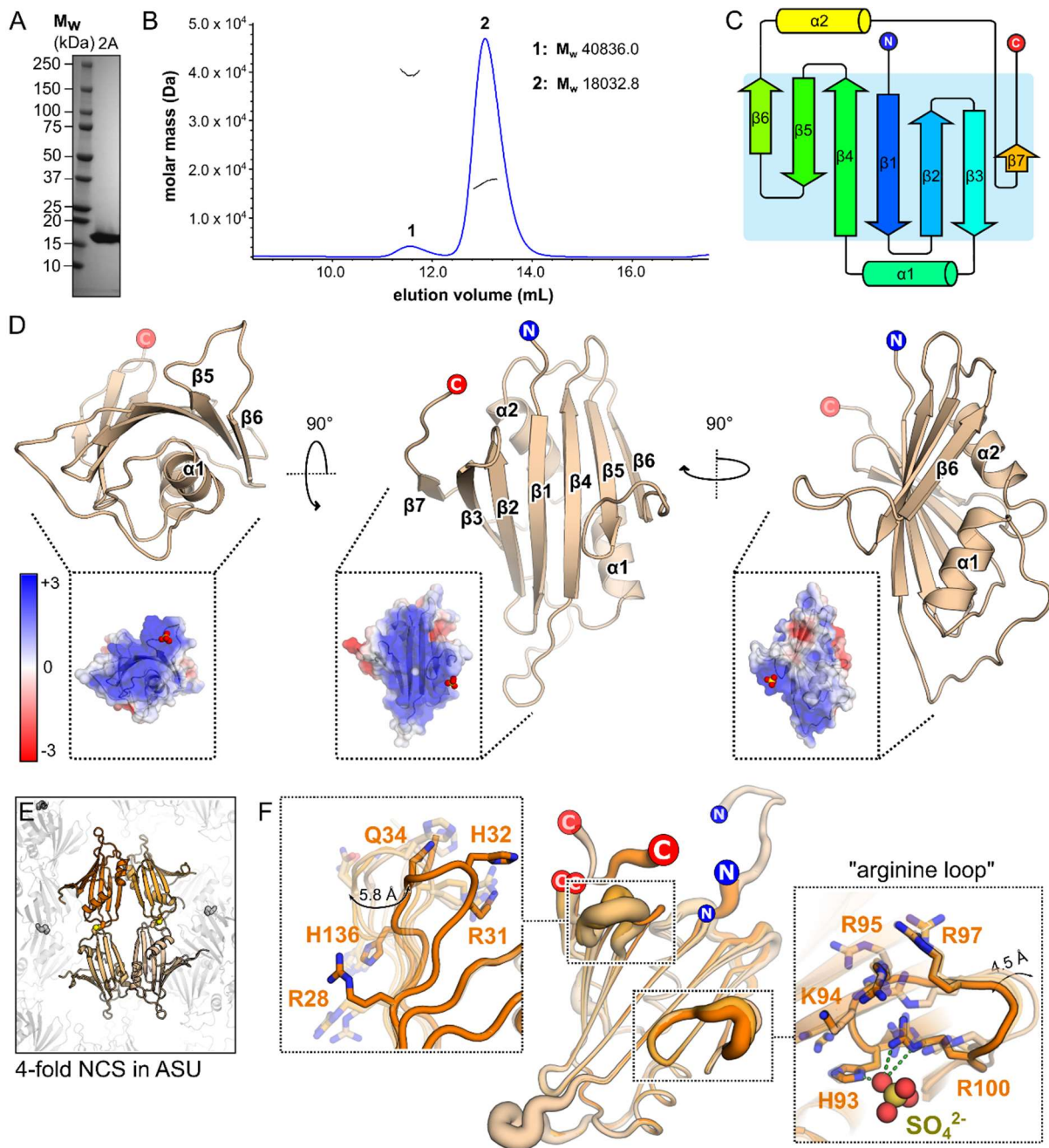
Acknowledgements

Figures and Legends

Figure 1. 2A adopts a highly basic RNA-binding fold with intrinsic flexibility. A) SDS-PAGE analysis of EMCV 2A after Ni-NTA, heparin affinity and size-exclusion chromatography. The gel was stained with Coomassie blue. **B)** SEC-MALS analysis of 5.2 mg/mL 2A in high-salt buffer. The differential refractive index is shown across the elution profile (blue) and weight-averaged molar masses of the indicated peaks are listed. **C)** Topological diagram of "beta-shell" fold: a curved central sheet comprising seven antiparallel beta strands, supported by two helices. **D)** X-ray crystal structure of EMCV 2A in three orthogonal views. N- and C- termini are indicated. *<Inset>* Electrostatic surface potential calculated at pH 7.4, coloured between +3 (blue) and -3 (red) kT/e⁻. **E)** Four molecules of 2A are present in the asymmetric unit of the crystal, arranged as two pairs of disulfide-linked dimers (spheres). **F)** Superposition of the four NCS-related 2A chains in E) reveals regions of conformational flexibility. The width of the cartoon is proportional to atomic B-factor. *<Insets>* Close-up view of surface loops exhibiting the greatest variation per molecule. Flexible sidechains are shown as sticks, and the Cα backbone deviation is indicated in Å. The position of a sulfate ion from the crystallisation buffer is indicated with spheres.
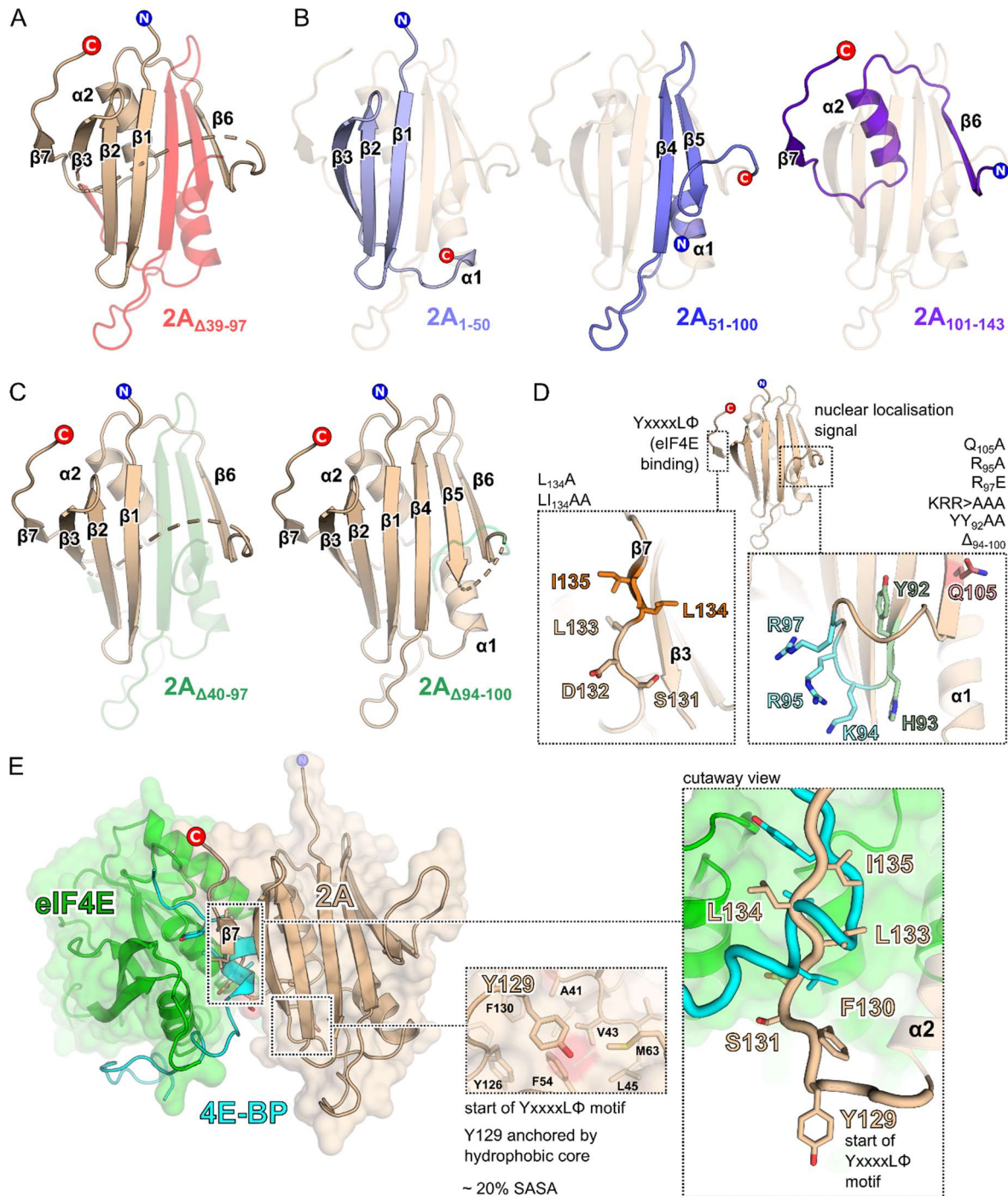
Supplementary Figure 1 - related to Figure 1. A) Structural consequences of the 2A$_{\Delta39-97}$ mutation described by Svitkin *et al.*[39] Deleted amino acids are highlighted in red. **B)** Truncation fragments 2A$_{1-50}$, 2A$_{51-100}$ and 2A$_{101-143}$ described by Petty *et al.*[40] In each case the remaining fragment is highlighted in blue and overlaid against the structure of the full protein for context. **C)** Deletion mutants 2A$_{\Delta40-97}$ and 2A$_{\Delta94-100}$ as described by Groppo *et al.*[25] Deleted amino acids are highlighted in green. **D)** Location of point-mutations made by Groppo *et al.*[25] in the putative nuclear localisation sequence and putative C-terminal YxxxxLΦ eIF4E binding motif. Mutated amino acids are shown as coloured sticks. **E)** Comparison 4E-BP1 YxxxxLΦ binding motif and the putative YxxxxLΦ motif in 2A. The crystal structure of the complex between eIF4E and 4E-BP1 is shown (green and blue, respectively) with 2A (wheat) docked via least-squares superposition of the YxxxxLΦ motif. *<Insets>* Contrast between the 2A YxxxxLΦ motif, in an extended β-strand conformation (wheat), and the 4E-BP1 YxxxxLΦ motif, in a compact α-helical conformation, with Y129 partially buried (~ 20% solvent-accessible residue surface area). 2A binding to eIF4E is not compatible with the known 4E-BP1 interface.
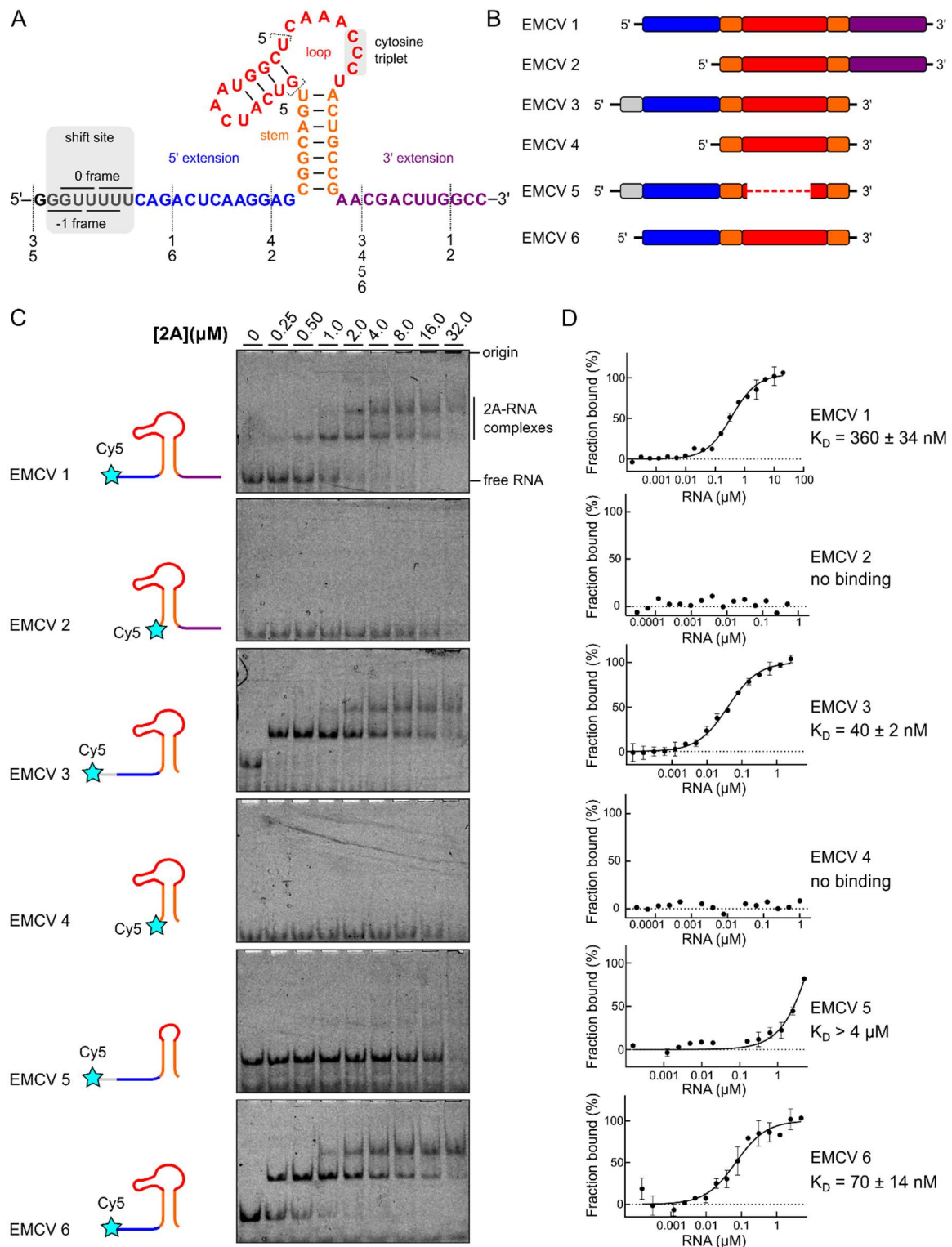
Figure 2. 2A binds to a minimal 47 nt element in the viral RNA. A and B) Sequences and schematic diagrams of the EMCV 1–6 constructs used to assay 2A binding. **C)** EMSA analyses showing that removal of the 5′ extension (blue) disables 2A binding. EMSAs were conducted with Cy5-labelled EMCV RNA (50 nM) at 2A concentrations varying between 0–32 µM. Following non-denaturing electrophoresis, fluorescence was imaged using a Typhoon scanner. **D)** Microscale thermophoresis (MST) was used to quantify the interactions observed in C). Binding affinities of unlabelled 2A to fluorescently labelled EMCV RNA (5 nM) were measured using Monolith NT.Pico (NanoTemper Technologies) at 5% LED power and medium MST power. All measurements were repeated twice. RNA concentration ranges of 60 pM – 20 µM (EMCV 1) and 150 pM – 5 µM (EMCV 2–6) were used.
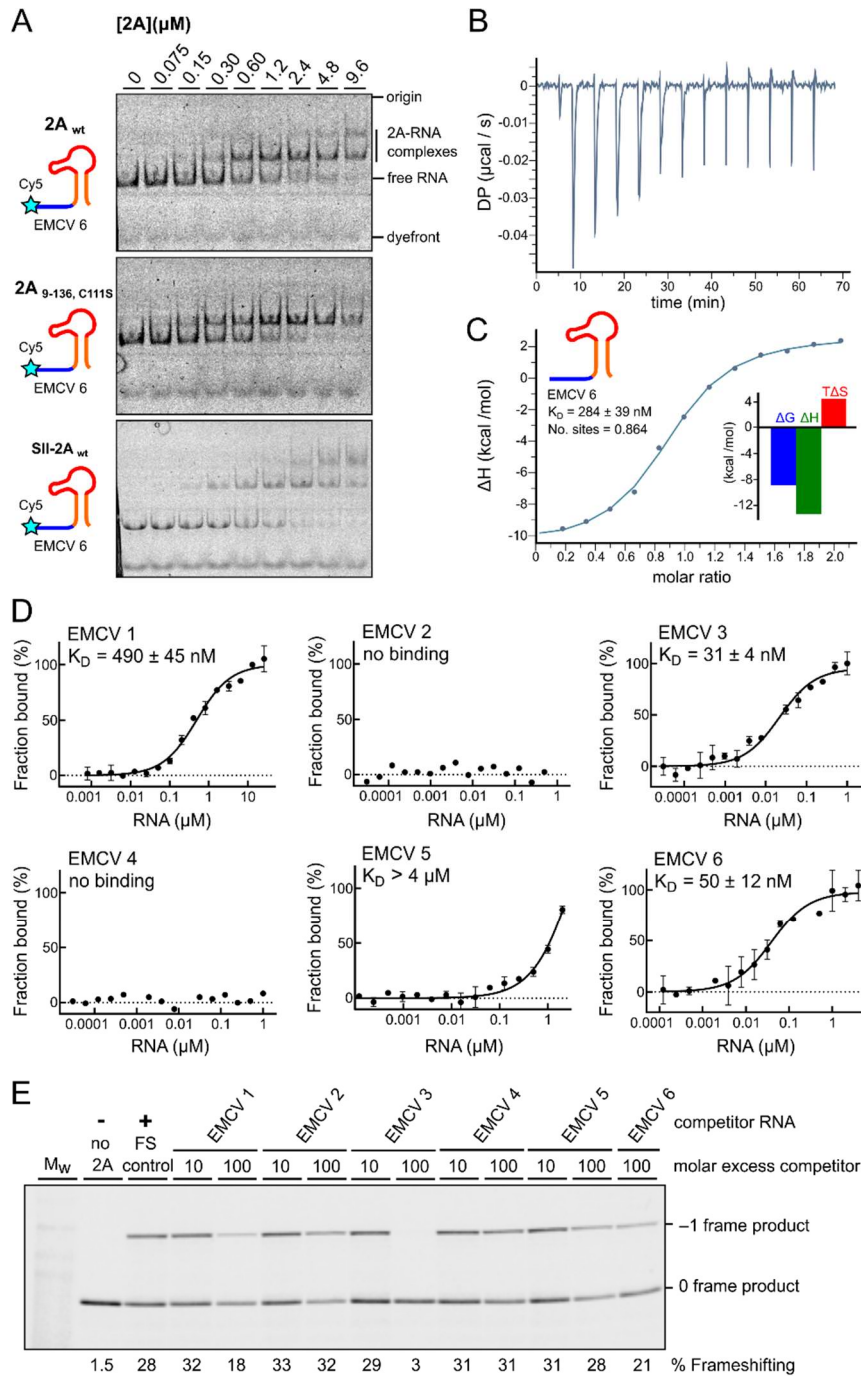
Supplementary Figure 2 - related to Figure 2. A) Side-by-side comparisons of 2A$_{wt}$, 2A$_{9-136; C111S}$ and SII-2A$_{wt}$. Equivalent RNA binding is observed in all cases by EMSA analyses conducted with 50 nM Cy5-labelled EMCV 6 RNA and 2A concentrations between zero and 9.6 µM. Following non-denaturing electrophoresis, fluorescence was imaged using a Typhoon scanner. **B)** Baseline-corrected differential power (DP) versus time for ITC titration of EMCV 6 RNA into 2A protein. **C)** Normalized binding curve showing integrated changes in enthalpy (ΔH) against molar ratio for titration in B), showing a ~1:1 molar ratio and nanomolar affinity *<Inset>* Histogram showing relative contributions of ΔH and TΔS terms to the overall exergonic interaction. **D)** MST binding curves and reported K$_D$ values of fluorescently labelled 2A protein (5 nM) and short unlabelled RNAs (as in Figure 2A and B) at concentrations between 800 pM – 26 µM for EMCV 1 and 120 pM – 4 µM for EMCV 2–6. **E)** Experiment showing the effects of titrating excess short RNAs (TMEV 1–6) as competitors into an *in vitro* frameshift reporter assay. The concentrations of the reporter mRNA and 2A were kept constant in the RRL and short RNAs were

added in 10- and 100- fold molar excess relative to the reporter mRNA, as indicated. Translation products were visualised by using ^{35}S-Met autoradiography, and % frameshifting was calculated following densitometry and correction for the number of methionines present in 0 frame and −1 frame products.
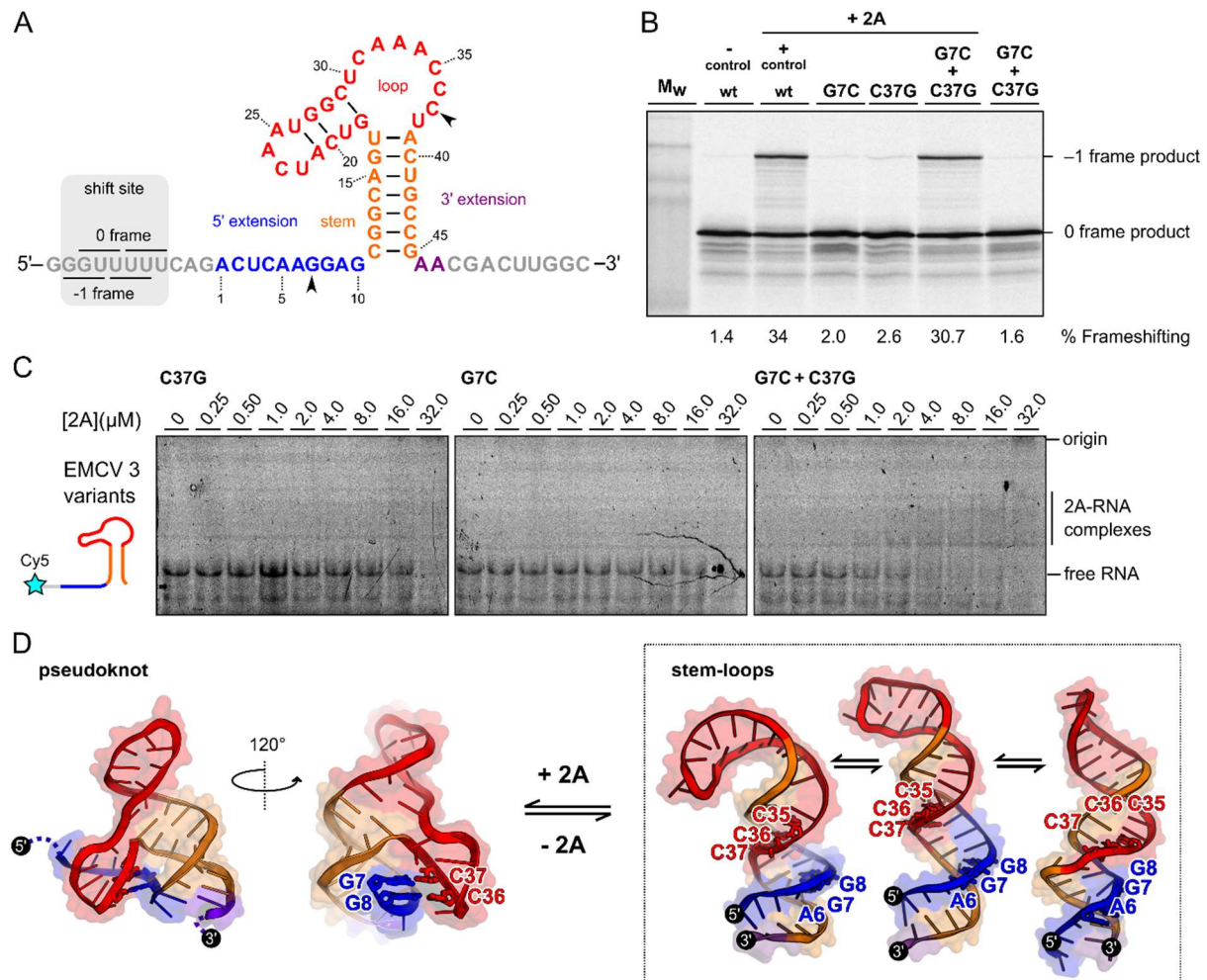
Supplementary Figure 3 - related to Figure 2. A) Schematic diagram showing numbered sequence of the EMCV 6 minimal PRF stimulatory element. **B)** Frameshifting assays showing evidence for a base-pairing interaction between G7 and C37. Individual G7C and C37C mutations reduce frameshifting to near-background levels. However, the double mutation (which would permit a compensatory C-G base-pair to form) restores frameshifting to wild-type levels. **C)** EMSA analyses showing that individual G7C and C37G mutations in the EMCV 6 RNA prevent 2A binding, but the double G7C+C37G mutation restores binding. Experiments were conducted with 50 nM Cy5-labelled EMCV 3 RNA variants and 2A concentrations between zero and 32 μM. Following non-denaturing electrophoresis, fluorescence was imaged using a Typhoon scanner. **D)** Equilibrium between several predicted stem-loops and alternate pseudoknot conformation, colour-coded as in A). Pseudoknot-like conformation involves base pairs between G7 and G8 in the 5′ extension and C36 and C37 in the loop (shown as sticks). These interactions are not maintained in any predicted stem-loop conformation.
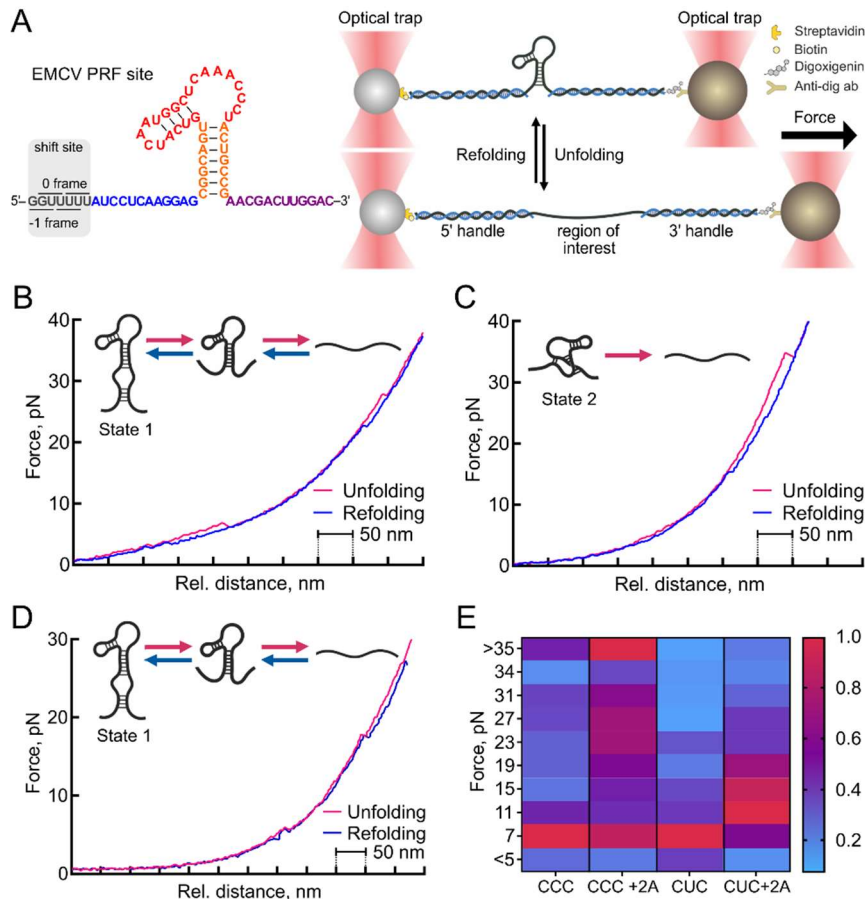
Figure 3. 2A increases resistance of the RNA element to mechanical unwinding. A) *<Left>* Primary sequence of the construct used in optical tweezer experiments, with predicted secondary structure, colour coded as in Figure 2. *<Right>* Schematic diagram illustrating the optical tweezer experiments (right). RNA is hybridized to ssDNA handles and immobilised on beads. These are used to exert pulling force on the RNA with a focused laser beam. **B)** Representative force-distance curve of the unfolding transition of the CCC (wild type) RNA element resulting from a stem-loop-like conformation of PRF site (state 1). The inferred (un)folding pathway is indicated at the upper-left corner of the graph. **C)** Representative force-distance curve of the unfolding transition of the CCC (wild type) RNA element resulting from a pseudoknot-like conformation of PRF site (state 2). The inferred unfolding pathway is indicated at the upper-left corner of the graph. **D)** Representative force-distance curve of the unfolding transition of the CUC (mutant) RNA element resulting from stem-loop-like (state 1) conformation of PRF site. The inferred (un)folding pathway is indicated at the upper-left corner of the graph. **E)** Heatmap showing normalized distribution of unfolding forces observed in the force spectroscopy experiments. We observe an increase in the higher force (~25 pN) and very high force (>35 pN) unfolding events for CCC RNA in the presence of 2A protein (CCC+2A) compared to the RNA-only sample (CCC). The CUC sample shows only minor changes in the presence of the EMCV 2A protein.
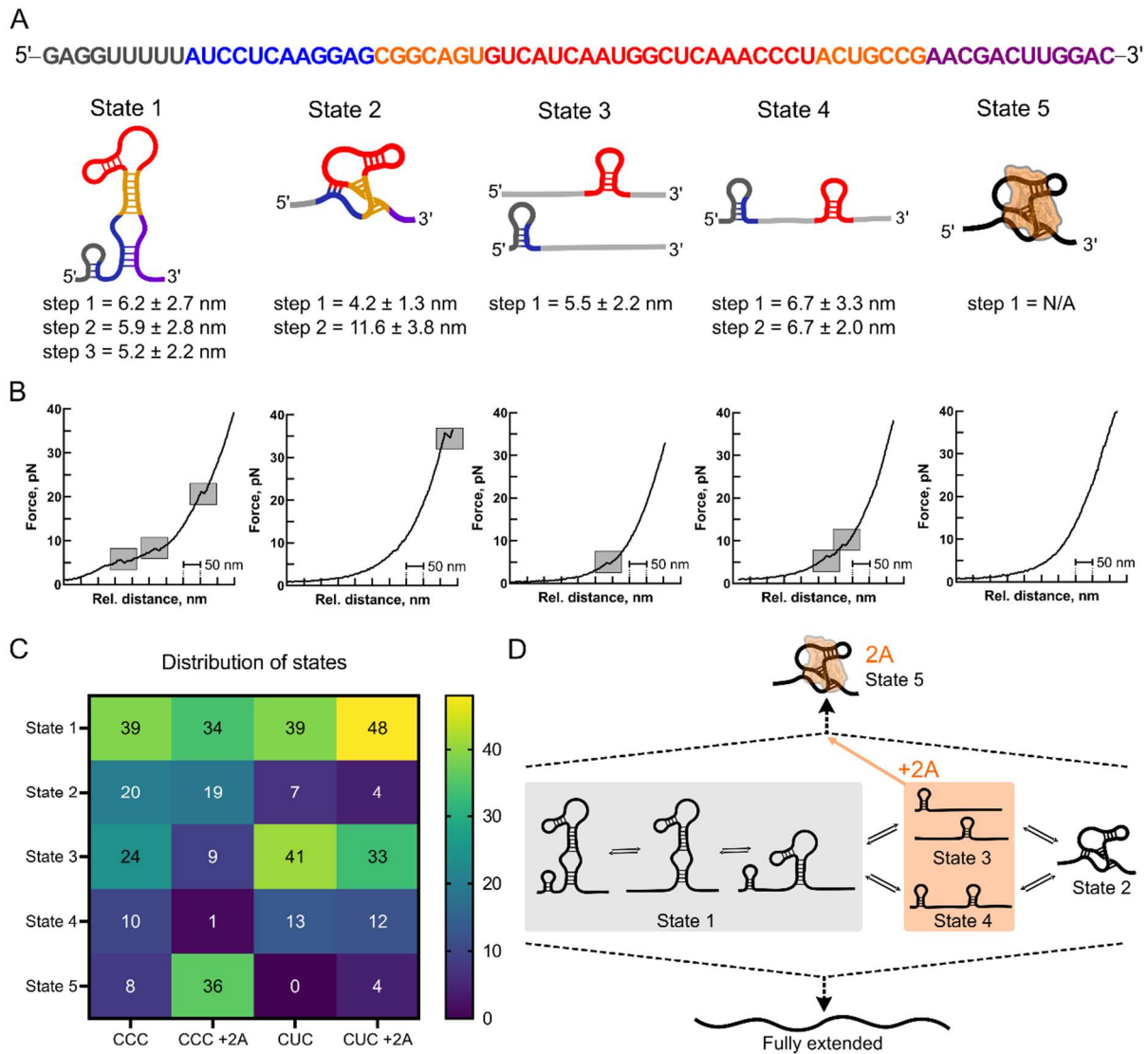
Figure 4. The EMCV RNA stimulatory element exists in several conformers. A) *<upper>* Primary sequence of the construct used in optical tweezer experiments, colour coded as in Figures 2 and 3. *<lower>* Conformational states observed during the measurements with proposed secondary structures. Step sizes correspond to the steps observed for the CCC (wild type) sample. **B)** Force-distance (FD) curves examples for each of the RNA conformers. Each graph represents the state depicted above the graph. Unfolding steps are marked with the grey squares. **C)** Heatmap showing population distribution of RNA conformers among the measured samples. In the presence of EMCV 2A protein, the CCC sample population profile shifts towards the more stable conformer (state 5) with a concomitant decrease in the population of the two low-force states (state 3 and 4). The CUC sample does not show a significant change in the population profile in the presence of EMCV 2A protein. **D)** Suggested model of the RNA conformation transitions. State 3 and 4 represent a partially folded state 1. State 1 alternates between the fully folded stated (showing a three-step FD profile) and two partially folded states (showing two-step FD profiles). State 2 is a stable pseudoknot conformation with high unfolding forces (~30 pN). State 5 is a stable conformer population that shows no unfolding steps under the measuring conditions. This state is more abundant in the presence of the EMCV 2A protein.
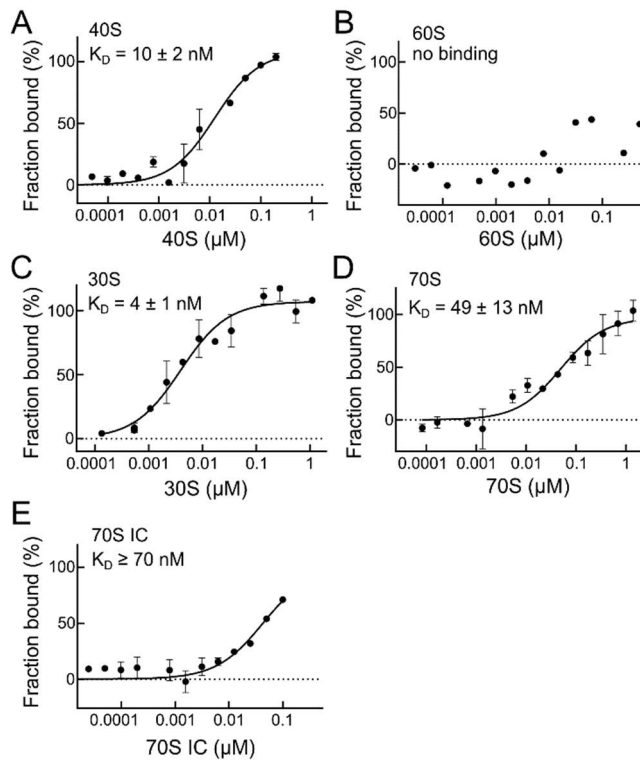
Figure 5. 2A binds directly to eukaryotic and prokaryotic ribosomes. In all MST experiments, fluorescently labelled 2A protein was used at a final concentration of 5.0 nM. **A and B)** Binding curves and reported K_D values using unlabelled 40S and 60S subunits at concentrations of 20 pM – 0.4 µM. 2A binds selectively to the small ribosomal subunit. **C)** Binding curve and reported K_D value for 2A-30S interaction, showing a strong interaction with the prokaryotic small subunit. 30S concentrations of 30 pM – 1 µM were used. **D and E)** Binding curves and reported K_D values for 2A-70S and 2A-70S$_{IC}$ (initiated complex) interactions, respectively. Concentrations of 800 pM – 1.375 µM (70S) and 3 pM - 0.1 µM for 70S IC (70S IC) were used.
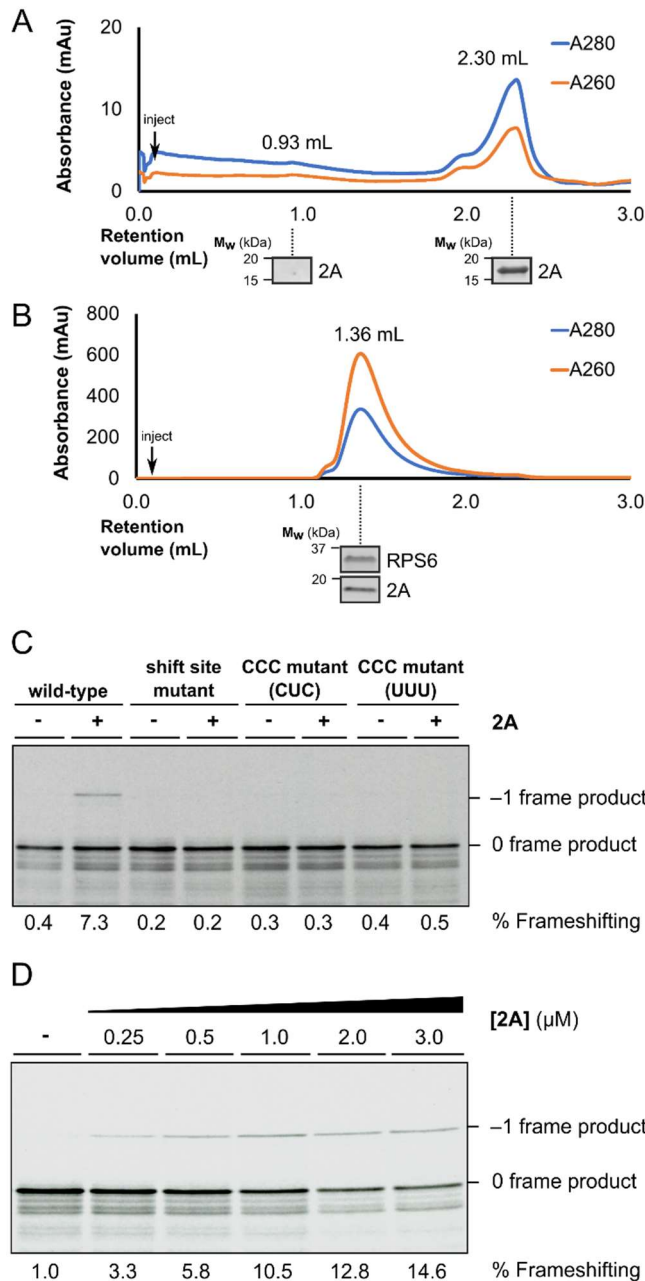
Supplementary Figure 4 - related to Figure 5. A) Size-exclusion chromatogram of purified 2A on a Superdex 6 3.2/300 column. 2A is only detectable by immunoblot in the late-eluting peak and not in early fractions. **B)** as in A), following the incubation of 2A with purified RRL 40S subunits. 2A co-migrates with the 40S peak as confirmed by immunoblot, indicative of binding. **C)** Frameshifting assay showing the reconstitution of 2A-dependent PRF in a prokaryotic *in vitro* translation system. Translation products were visualised by ^{35}S-Met autoradiography and % frameshifting was calculated following densitometry and correction for the number of methionines present in 0 frame and −1 frame products. **D)** PRF efficiency in the prokaryotic system is proportional to 2A concentration. At high levels, 2A displays inhibitory effects on total translation. Data analysed as above.
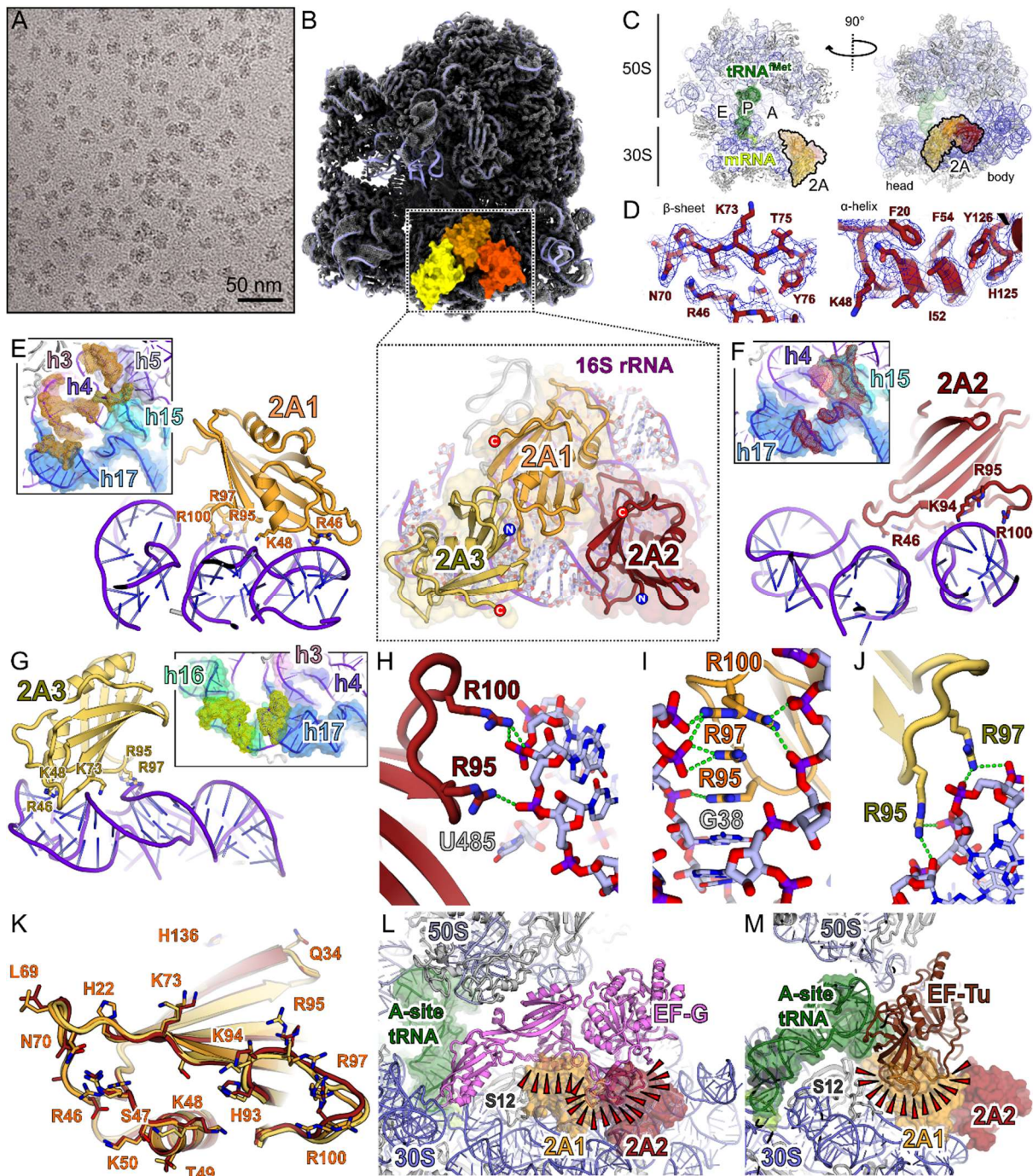
Figure 6. 2A binds to the 70S ribosome via interactions with the 16S rRNA. A) Cryo-EM analysis of a complex formed between initiated *E. coli* 70S ribosomes and EMCV 2A. Images (× 75,000, total dose 54.4 e⁻ / Å^2) were recorded on a Titan Krios microscope (300 kV, Falcon III detector). **B)** Cryo-EM electron density map at 2.6 Å resolution after focus-classification and refinement. Three copies of 2A (orange, red, yellow) are bound to the 16S rRNA of the small (30S) subunit. **C)** Ribbon diagram of initiated 70S-mRNA-tRNA^{fMet}-2A complex. Ribosome sites are labelled A, P and E. The initiator tRNA^{fMet} (dark green), mRNA (light green), and 2A (orange, red, yellow) are shown in two orthogonal views. **D)** Examples of local density for 2A. Well-resolved sidechains are clearly visible in both beta strands and alpha helices. **E – G)** Details of 2A interaction with 16S rRNA (purple). Residues involved in interactions are labelled and shown as sticks <*Inset*> View of the rRNA surface bound by each copy of 2A. The rRNA helices are colour-coded and labelled. The 2A contact surface is shown as a coloured mesh (orange, red and yellow, respectively). **H – J)** Details of interactions between 2A R95, R97 and R100 (sticks) and the rRNA backbone (sticks) for each copy of 2A (orange, red, yellow). Polar or electrostatic

contacts are indicated by a green dashed line. **K)** Superposition of the three copies of 2A to highlight conformational flexibility. Residues involved in rRNA binding are labelled and shown as sticks. **L)** Comparison of 70S-2A complex to 70S pre-translocation complex with EF-G (4V7D). 2A binding would clash (red wedges) with EF-G binding. **M)** Comparison of 70S-2A complex to 70S complex with EF-Tu (5WE6). 2A binding would clash (red wedges) with EF-Tu binding.
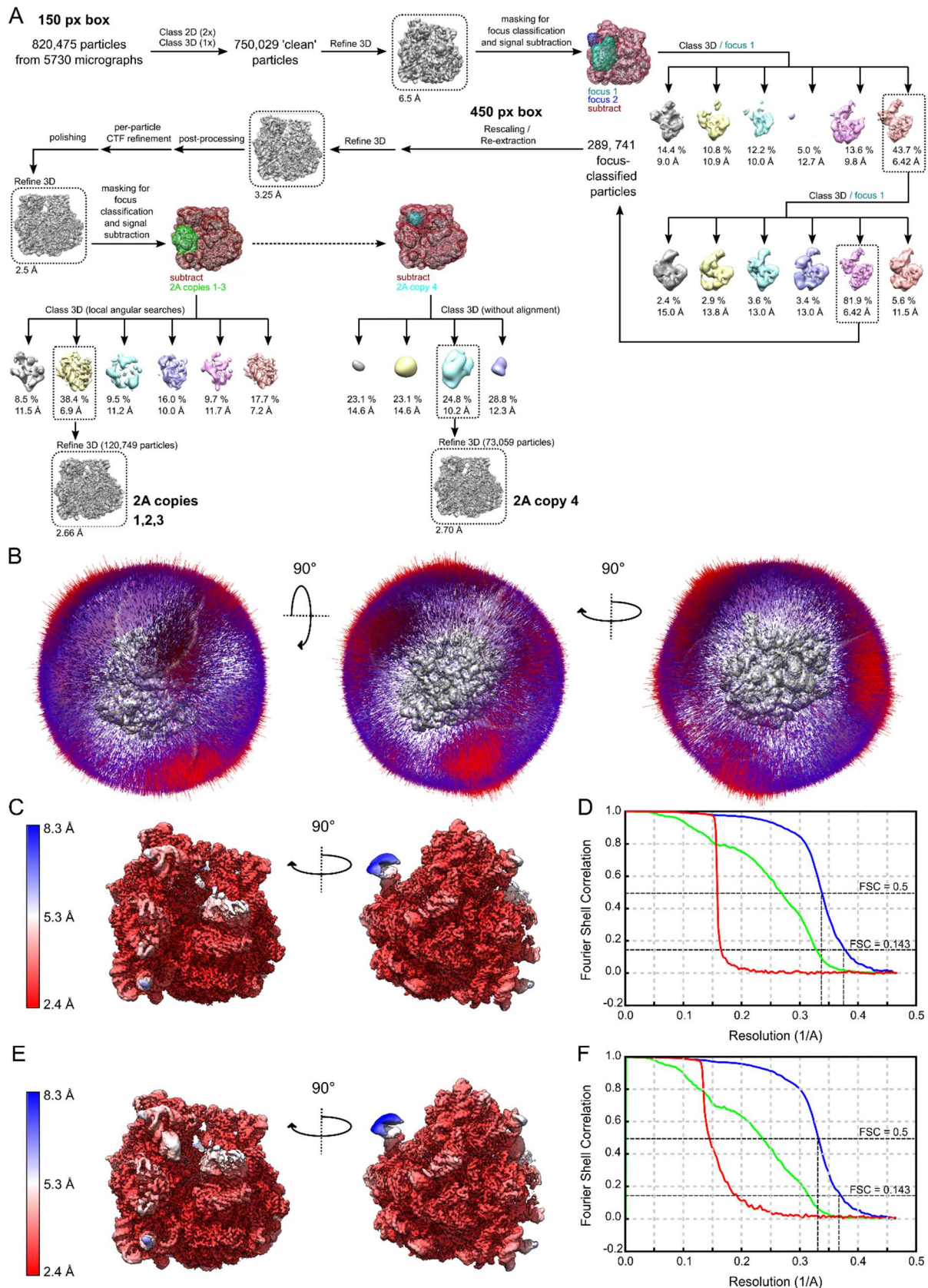
Supplementary Figure 5 - related to Figure 6. A) Schematic summary of steps in cryo-EM data processing. **B)** Three orthogonal views showing the angular distribution of particles contributing to the final 3D reconstruction. This is shown for the highest-resolution Refine3D result i.e. immediately after particle polishing. **C)** Local-resolution map for the final reconstruction of 70S-2A₃. The surface is

coloured by local resolution from red (highest; 2.4 Å) to blue (lowest; 8.3 Å). **D)** Gold-standard Fourier shell correlation (FSC) curve for the 70S-2A$_3$ map. Masked (blue), unmasked (green) and phase-randomised masked (red) plots are shown. **E)** Local-resolution map for the final reconstruction of 70S-2A$_4$, details as in C). **F)** Gold-standard Fourier shell correlation (FSC) curve for the 70S-2A$_4$ map. Details as in D).
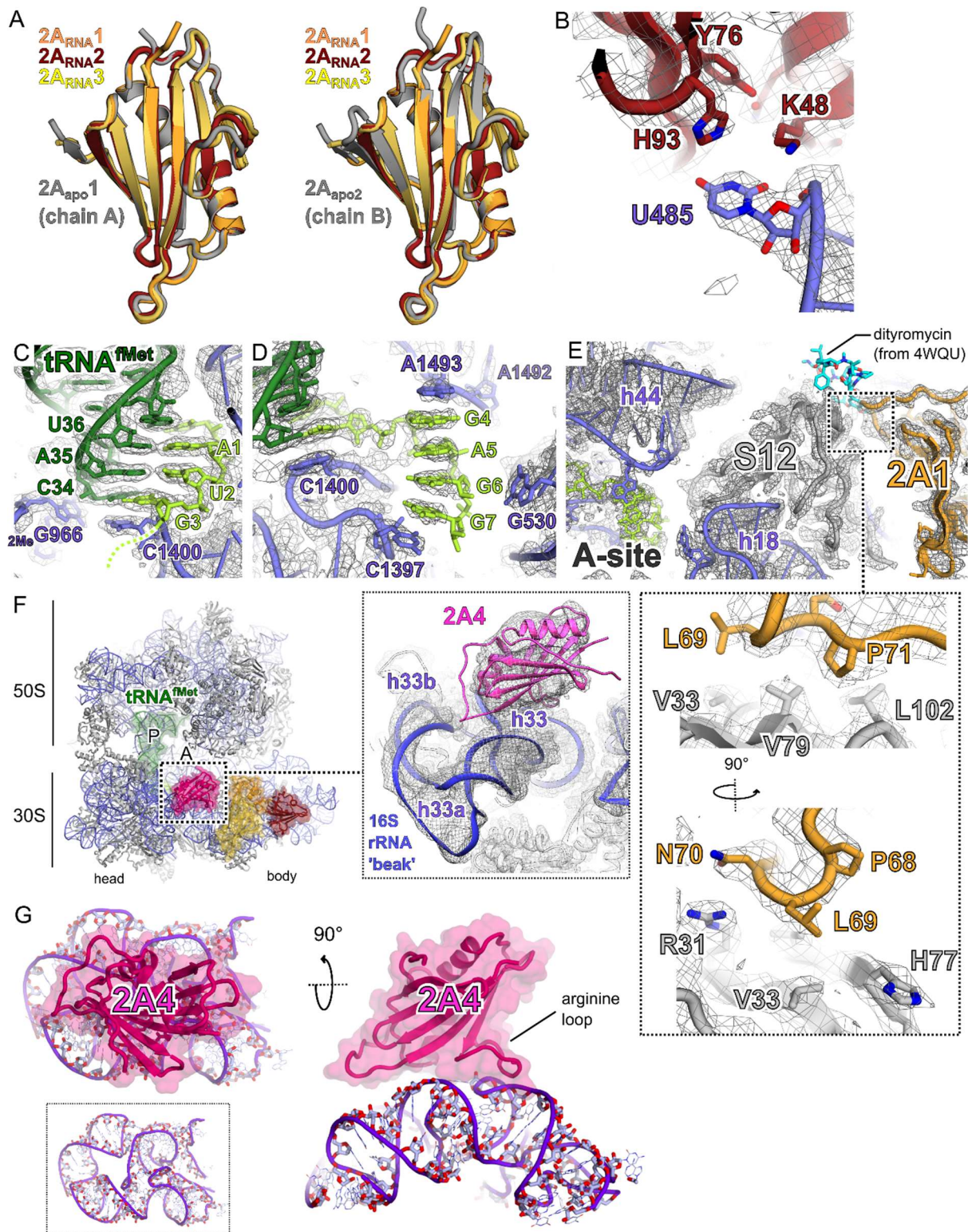
Supplementary Figure 6 - related to Figure 6. A) Comparison between conformations of 2A protein in RNA bound states (orange, red, yellow) and the two unliganded states observed by NCS in the crystal structure. The $2A_{apo}1$ conformation observed in chain A is most similar to the RNA-bound state. Structural alignments were performed by least-squares superposition of the C_α backbone. **B)** Details of a base-specific interaction between U485 (helix 17 of 16S) and a pocket on the surface of 2A2 (red). **C)** Cryo-EM density at the P-site. Codon-anticodon pairing between the mRNA (lime) and the initiator tRNA^{fMet} (dark green). The tRNA is in an undistorted P/P conformation as expected. **D)** Cryo-EM density at the A-site, coloured as in B). Additional 30S residues with roles in decoding are shown as sticks

(purple). **E)** Details of a hydrophobic 2A1 interaction with ribosomal protein S12. The contact surface is on the factor-binding face of S12, away from the decoding centre. The binding site of antibiotic dityromycin on S12 (from 4WQU) is shown with blue sticks. **F)** Ribbon diagram of initiated 70S-mRNA-tRNA^{fMet}-2A complex showing the location of the fourth copy of 2A (pink) present in a smaller population of particles. Ribosome sites are labelled A and P. The initiator tRNA^{fMet} (dark green), mRNA (lime), 2A1 (orange), 2A2 (red) and 2A3 (yellow) are also shown. *<Inset>* Section of the 70S-2A$_4$ local resolution map showing electron density at the 2A4 binding site. 2A4 binds to the 3′ major 'beak' domain of the 16S rRNA present in the 30S 'head', via electrostatic interactions with the ribose phosphate backbone of helices 33 and 34. **G)** Details of 2A4 interaction with 16S rRNA (purple) in two orthogonal views.
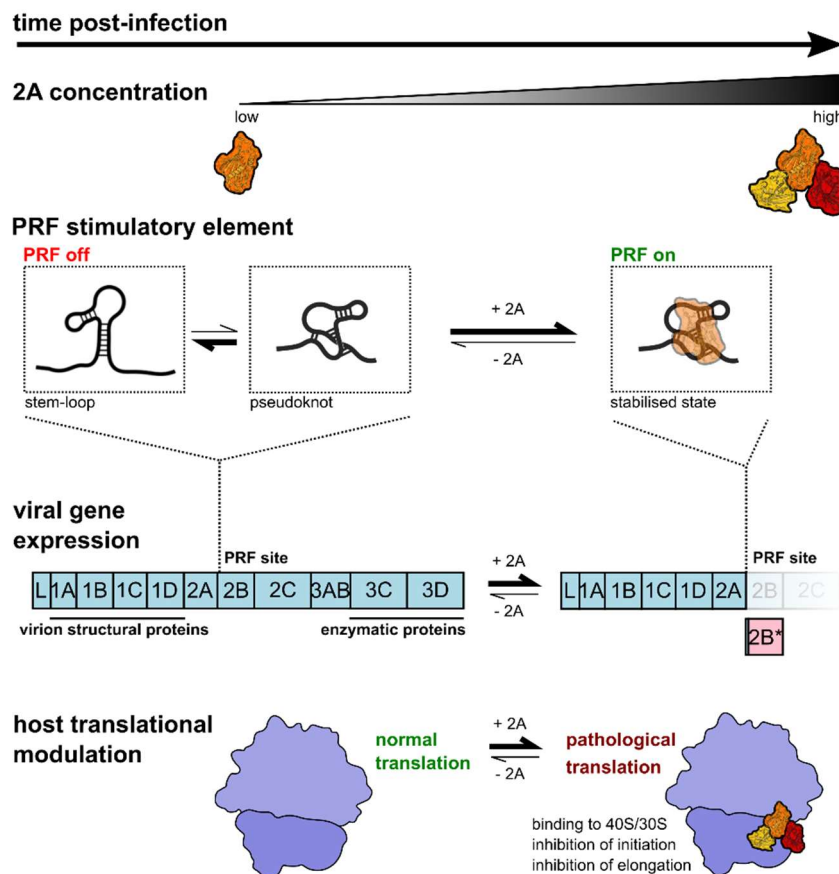
Figure 7. Molecular basis for 2A-induced reprogramming of gene expression. As 2A accumulates during EMCV infection, it selectively binds to and stabilises a pseudoknot-like conformation of the PRF stimulatory element, thereby enabling PRF, producing *trans*-frame product 2B* and downregulating the expression of enzymatic viral proteins later in infection. 2A also binds directly to the small ribosomal subunit at the translational GTPase factor binding site, progressively inhibiting both initiation and elongation as it accumulates. This may contribute to the shutdown of host cell translation during lytic infection.

Table 1 – Crystallographic data collection and refinement. Data were recorded from a single native crystal and a single crystal of selenomethionine-derivatised protein. Values for high resolution shells are shown in parentheses.

	Native	Selenomethionine derivative		
		Remote (High E)	Peak	Inflection
Data collection				
Wavelength (Å)	0.97958	0.97635	0.97965	0.97974
Space group	$P6_222$		$P6_222$	
Cell dimensions				
a, b, c (Å)	91.54, 91.54, 316.50	91.17, 91.17, 316.75	91.07, 91.07, 316.31	91.42, 91.42, 317.43
Resolution (Å)	43.91–2.62 (2.67–2.62)	79.0–2.98 (3.03–2.98)	76.53–2.91 (2.96–2.91)	79.36–3.08 (3.13–3.08)
Unique reflections	24,668 (1179)	16,809 (817)	17,967 (869)	15,435 (741)
Completeness (%)	100.0 (99.8)	99.6 (99.0)	99.7 (99.2)	99.9 (99.6)
Anomalous	-	99.1 (98.7)	99.2 (98.8)	99.3 (99.2)
Multiplicity	18.9 (19.4)	37.4 (38.5)	18.6 (17.2)	18.6 (19.8)
Anomalous		21.0 (20.9)	10.4 (9.3)	10.5 (10.7)
R_{merge}	0.227 (2.766)	0.293 (2.929)	0.262 (2.113)	0.284 (2.398)
R_{pim}	0.053 (0.640)	0.048 (0.473)	0.062 (0.521)	0.067 (0.548)
$CC_{1/2}$	0.998 (0.728)	0.999 (0.834)	0.997 (0.804)	0.997 (0.822)
CC_{anom}	-	0.314 (0.002)	0.400 (0.024)	0.167 (0.000)
Mean $I/\sigma(I)$	11.5 (1.0)	12.9 (1.2)	9.9 (1.2)	9.3 (1.1)
Wilson B ($Å^2$)	54.9			
Refinement				
Resolution (Å)	43.91–2.62 (2.71–2.62)			
Reflections				
Working set	23,333 (2223)			
Test set	1259 (139)			
R_{work}	0.2080			
R_{free}	0.2535			
No. of atoms				
Protein	4414			
Solvent	114			
Other*	15			
Root mean square deviation				
Bond lengths (Å)	0.007			
Bond angles (°)	1.077			
Ramachandran favoured (%)	94.95			
Ramachandran outliers (%)	0.97			
Poor rotamers (%)	1.46%			
Mean B value ($Å^2$)	71.2			

*sulfate ions

Table 2 - Summary of dissociation constants (K_D) measured by microscale thermophoresis with various 2A interaction partners. ND: Not determined.

	K_D (LABELLED 2A)	K_D (LABELLED RNA)
EMCV 1	360 ± 34 nM	490 ± 45 nM
EMCV 2	No binding	No binding
EMCV 3	40 ± 2 nM	31 ± 4 nM
EMCV 4	No binding	No binding
EMCV 5	> 4 µM (does not reach saturation)	> 4 µM (does not reach saturation)
EMCV 6	70 ± 14 nM	50 ± 12 nM
40S	10 ± 2 nM	ND
60S	No binding	ND
30S	4 ± 1 nM	ND
70S	49 ± 13 nM	ND
70S IC	> 60 nM (does not reach saturation)	ND

Table 3 - Averaged unfolding force and extension values from the optical tweezer measurements on the wild-type CCC and loop mutant CUC RNA in the presence and absence of the 2A protein.

			CCC	CCC + 2A	CUC	CUC + 2A
State 1	step 0	F (pN)	6.4±1.8	5.9±1.3	5.4±1.3	7.2±2.6
		Δx (nm)	6.2±2.7	5.2±1.4	5.4±1.7	5.1±1.1
	step 1	F (pN)	10.4±4.2	13.5±5.2	9.1±4.4	13.3±3.7
		Δx (nm)	5.9±2.8	5.2±1.9	5.4±1.8	6.8±2.8
	step 2	F (pN)	25.4±6.3	26.3±6.8	20.4±5.9	25.7±7.4
		Δx (nm)	5.2±2.2	5.4±1.6	5.6±2.1	5.1±1.9
State 2	step 0	F (pN)	8.5±5.3		9.7±0	
		Δx (nm)	4.2±1.3		5.1±0	
	step 1	F (pN)	29.2±7.9	30.9±6.5	31.1±10.6	21.1±0.4
		Δx (nm)	11.6±3.8	10.4±2.5	10.8±0.9	9.5±1.2
State 3	step 0	F (pN)	7.0±2.3	8.8±3.1	6.4±2.7	13.6±4.8
		Δx (nm)	5.5±2.2	7.8±3.5	7.0±4.0	6.9±1.9
State 4	step 0	F (pN)	5.3±0.9	5.2±0	5.6±1.0	8.7±1.8
		Δx (nm)	6.7±3.3	3.8±0	4.9±1.3	5.8±1.0
	step 1	F (pN)	8.8±1.6	9.8±0	8.9±2.0	11.6±1.3
		Δx (nm)	6.7±2.0	7.0±0	5.6±1.9	8.1±1.4
State 5	No step	F (pN)	>35	>35	>35	>35
		Δx (nm)	NA	NA	NA	NA

Table 4 – Cryo-EM data collection, processing and refinement

Data collection	
Microscope	Titan Krios
Detector	Falcon III (integration)
Magnification	75,000x
C2 aperture (μm)	50
Objective aperture (μm)	100
Pixel size (Å)	1.07
Voltage (kV)	300
Electron dose (e$^-$/Å^2)	54.4
Defocus range (μm)	−1.2, −1.5, −1.8, −2.1, −2.4, −2.7, −3.0
Phase shift range	-
Number of micrographs	5730
Processing	
No. particles	120,749
Resolution (FSC 0.143)	2.66
Model	
Composition (#)	
Chains	59
Atoms	149908
Residues	protein: 6354
	nucleic acid: 4641
Water	2
Ligands	Zn: 2
	Mg: 437
Bond RMSD	
Length (Å) (# > 4s)	0.013 (4)
Angles (°) (# > 4s)	0.955 (36)
MolProbity score	2.41
Clash score	6.53
Ramachandran plot (%)	
Outliers	0.08
Allowed	4.84
Favored	95.08
Rotamer outliers (%)	8.65
Cß outliers (%)	0.00
Cis proline/general	1.4/0.0
Twisted proline/general	0.0/0.1
CaBLAM outliers (%)	2.28
ADP (B-factors)	
Iso/Aniso (#)	149908/0
min/max/mean	
protein	14.04/115.80/54.21
nucleotide	8.93/160.74/49.96
ligand	15.68/116.89/35.10
water	18.55/34.92/26.73

Occupancy	
Mean	1.00
occ = 1 (%)	100.0
0 < occ < 1 (%)	0.0
occ > 1 (%)	0.0
Data	
Box	
Lengths (Å)	241.82, 257.87, 273.92
Angles (°)	90.00, 90.00, 90.00
Supplied Resolution (Å)	2.4
Resolution Estimates (Å)	
d 99 (full))	2.6
d model	2.6
d FSC model (0/0.143/0.5)	2.3/2.3/2.6
Map min/max/mean	-0.50/1.01/0.00
Model vs. Data	
CC (mask)	0.88
CC (box)	0.85
CC (peaks)	0.83
CC (volume)	0.87
Mean CC for ligands	0.68

References

1   Jang, S. K. *et al.* A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. *J Virol* **62**, 2636-2643 (1988).

2   Palmenberg, A. C. *et al.* Proteolytic processing of the cardioviral P2 region: primary 2A/2B cleavage in clone-derived precursors. *Virology* **190**, 754-762, doi:10.1016/0042-6822(92)90913-a (1992).

3   Hahn, H. & Palmenberg, A. C. Deletion mapping of the encephalomyocarditis virus primary cleavage site. *J Virol* **75**, 7215-7218, doi:10.1128/JVI.75.15.7215-7218.2001 (2001).

4   Napthine, S. *et al.* Protein-directed ribosomal frameshifting temporally regulates gene expression. *Nat Commun* **8**, 15582, doi:10.1038/ncomms15582 (2017).

5   Loughran, G., Firth, A. E. & Atkins, J. F. Ribosomal frameshifting into an overlapping gene in the 2B-encoding region of the cardiovirus genome. *Proc Natl Acad Sci U S A* **108**, E1111-1119, doi:10.1073/pnas.1102932108 (2011).

6   Firth, A. E. & Brierley, I. Non-canonical translation in RNA viruses. *The Journal of general virology* **93**, 1385-1409, doi:10.1099/vir.0.042499-0 (2012).

7   Atkins, J. F., Loughran, G., Bhatt, P. R., Firth, A. E. & Baranov, P. V. Ribosomal frameshifting and transcriptional slippage: From genetic steganography and cryptography to adventitious use. *Nucleic Acids Res* **44**, 7007-7078, doi:10.1093/nar/gkw530 (2016).

8   Korniy, N., Samatova, E., Anokhina, M. M., Peske, F. & Rodnina, M. V. Mechanisms and biomedical implications of -1 programmed ribosome frameshifting on viral and bacterial mRNAs. *FEBS Lett* **593**, 1468-1482, doi:10.1002/1873-3468.13478 (2019).

9   Chen, J. *et al.* Dynamic pathways of -1 translational frameshifting. *Nature* **512**, 328-332, doi:10.1038/nature13428 (2014).

10   Caliskan, N., Katunin, V. I., Belardinelli, R., Peske, F. & Rodnina, M. V. Programmed -1 frameshifting by kinetic partitioning during impeded translocation. *Cell* **157**, 1619-1631, doi:10.1016/j.cell.2014.04.041 (2014).

11   Choi, J., O'Loughlin, S., Atkins, J. F. & Puglisi, J. D. The energy landscape of -1 ribosomal frameshifting. *Sci Adv* **6**, eaax6969, doi:10.1126/sciadv.aax6969 (2020).

12   Namy, O., Moran, S. J., Stuart, D. I., Gilbert, R. J. & Brierley, I. A mechanical explanation of RNA pseudoknot function in programmed ribosomal frameshifting. *Nature* **441**, 244-247, doi:10.1038/nature04735 (2006).

13   Giedroc, D. P. & Cornish, P. V. Frameshifting RNA pseudoknots: structure and mechanism. *Virus Res* **139**, 193-208, doi:10.1016/j.virusres.2008.06.008 (2009).

14   Chen, G., Chang, K. Y., Chou, M. Y., Bustamante, C. & Tinoco, I., Jr. Triplex structures in an RNA pseudoknot enhance mechanical stability and increase efficiency of -1 ribosomal frameshifting. *Proc Natl Acad Sci U S A* **106**, 12706-12711, doi:10.1073/pnas.0905046106 (2009).

15   Ritchie, D. B., Foster, D. A. & Woodside, M. T. Programmed -1 frameshifting efficiency correlates with RNA pseudoknot conformational plasticity, not resistance to mechanical unfolding. *Proc Natl Acad Sci U S A* **109**, 16167-16172, doi:10.1073/pnas.1204114109 (2012).

16   Halma, M. T. J., Ritchie, D. B., Cappellano, T. R., Neupane, K. & Woodside, M. T. Complex dynamics under tension in a high-efficiency frameshift stimulatory structure. *Proc Natl Acad Sci U S A* **116**, 19500-19505, doi:10.1073/pnas.1905258116 (2019).

17   Ritchie, D. B., Soong, J., Sikkema, W. K. & Woodside, M. T. Anti-frameshifting ligand reduces the conformational plasticity of the SARS virus pseudoknot. *J Am Chem Soc* **136**, 2196-2199, doi:10.1021/ja410344b (2014).

18   Jackson, R. J. A detailed kinetic analysis of the in vitro synthesis and processing of encephalomyocarditis virus products. *Virology* **149**, 114-127, doi:10.1016/0042-6822(86)90092-9 (1986).

19   Yang, X. *et al.* Structures and Corresponding Functions of Five Types of Picornaviral 2A Proteins. *Front Microbiol* **8**, 1373, doi:10.3389/fmicb.2017.01373 (2017).

20   Aminev, A. G., Amineva, S. P. & Palmenberg, A. C. Encephalomyocarditis viral protein 2A localizes to nucleoli and inhibits cap-dependent mRNA translation. *Virus Res* **95**, 45-57, doi:10.1016/s0168-1702(03)00162-x (2003).

21   Aminev, A. G., Amineva, S. P. & Palmenberg, A. C. Encephalomyocarditis virus (EMCV) proteins 2A and 3BCD localize to nuclei and inhibit cellular mRNA transcription but not rRNA transcription. *Virus Res* **95**, 59-73, doi:10.1016/s0168-1702(03)00163-1 (2003).

22   Groppo, R. & Palmenberg, A. C. Cardiovirus 2A protein associates with 40S but not 80S ribosome subunits during infection. *J Virol* **81**, 13067-13074, doi:10.1128/JVI.00185-07 (2007).

23   Carocci, M. *et al.* Encephalomyocarditis virus 2A protein is required for viral pathogenesis and inhibition of apoptosis. *J Virol* **85**, 10741-10754, doi:10.1128/JVI.00394-11 (2011).

24   Mosenkis, J. *et al.* Shutoff of host translation by encephalomyocarditis virus infection does not involve cleavage of the eucaryotic initiation factor 4F polypeptide that accompanies poliovirus infection. *J Virol* **54**, 643-645 (1985).

25   Groppo, R., Brown, B. A. & Palmenberg, A. C. Mutational analysis of the EMCV 2A protein identifies a nuclear localization signal and an eIF4E binding site. *Virology* **410**, 257-267, doi:10.1016/j.virol.2010.11.002 (2011).

26   Merrick, W. C. eIF4F: a retrospective. *J Biol Chem* **290**, 24091-24099, doi:10.1074/jbc.R115.675280 (2015).

27   Visscher, K. -1 Programmed Ribosomal Frameshifting as a Force-Dependent Process. *Prog Mol Biol Transl Sci* **139**, 45-72, doi:10.1016/bs.pmbts.2015.11.003 (2016).

28   Mandal, S., Hoque, M. E. & Mao, H. Single-Molecule Investigations of G-Quadruplex. *Methods in molecular biology (Clifton, N.J.)* **2035**, 275-298, doi:10.1007/978-1-4939-9666-7_16 (2019).

29   Zhong, Z. *et al.* Mechanical unfolding kinetics of the SRV-1 gag-pro mRNA pseudoknot: possible implications for -1 ribosomal frameshifting stimulation. *Sci Rep* **6**, 39549, doi:10.1038/srep39549 (2016).

30   Yang, L. *et al.* Single-Molecule Mechanical Folding and Unfolding of RNA Hairpins: Effects of Single A-U to A.C Pair Substitutions and Single Proton Binding and Implications for mRNA Structure-Induced -1 Ribosomal Frameshifting. *J Am Chem Soc* **140**, 8172-8184, doi:10.1021/jacs.8b02970 (2018).

31   Chandra, V., Hannan, Z., Xu, H. & Mandal, M. Single-molecule analysis reveals multi-state folding of a guanine riboswitch. *Nat Chem Biol* **13**, 194-201, doi:10.1038/nchembio.2252 (2017).

32   Desai, V. P. *et al.* Co-temporal Force and Fluorescence Measurements Reveal a Ribosomal Gear Shift Mechanism of Translation Regulation by Structured mRNAs. *Mol Cell* **75**, 1007-1019 e1005, doi:10.1016/j.molcel.2019.07.024 (2019).

33   Wen, J. D. *et al.* Following translation by single ribosomes one codon at a time. *Nature* **452**, 598-603, doi:10.1038/nature06716 (2008).

34   Green, L., Kim, C. H., Bustamante, C. & Tinoco, I., Jr. Characterization of the mechanical unfolding of RNA pseudoknots. *J Mol Biol* **375**, 511-528, doi:10.1016/j.jmb.2007.05.058 (2008).

35   Chen, Y. T. *et al.* Coordination among tertiary base pairs results in an efficient frameshift-stimulating RNA pseudoknot. *Nucleic Acids Res* **45**, 6011-6022, doi:10.1093/nar/gkx134 (2017).

36   Yan, S., Wen, J. D., Bustamante, C. & Tinoco, I., Jr. Ribosome excursions during mRNA translocation mediate broad branching of frameshift pathways. *Cell* **160**, 870-881, doi:10.1016/j.cell.2015.02.003 (2015).

37   Niesen, F. H., Berglund, H. & Vedadi, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protoc* **2**, 2212-2221, doi:10.1038/nprot.2007.321 (2007).

38   Moreira, I. S., Fernandes, P. A. & Ramos, M. J. Hot spots--a review of the protein-protein interface determinant amino-acid residues. *Proteins* **68**, 803-812, doi:10.1002/prot.21396 (2007).

39   Svitkin, Y. V., Hahn, H., Gingras, A. C., Palmenberg, A. C. & Sonenberg, N. Rapamycin and wortmannin enhance replication of a defective encephalomyocarditis virus. *J Virol* **72**, 5811-5819 (1998).

40   Petty, R. V., Basta, H. A., Bacot-Davis, V. R., Brown, B. A. & Palmenberg, A. C. Binding interactions between the encephalomyocarditis virus leader and protein 2A. *J Virol* **88**, 13503-13509, doi:10.1128/JVI.02148-14 (2014).

41   Magnus, M., Boniecki, M. J., Dawson, W. & Bujnicki, J. M. SimRNAweb: a web server for RNA 3D structure modeling with optional restraints. *Nucleic Acids Res* **44**, W315-319, doi:10.1093/nar/gkw279 (2016).

42   Ren, J., Rastegari, B., Condon, A. & Hoos, H. H. HotKnots: heuristic prediction of RNA secondary structures including pseudoknots. *RNA (New York, N.Y.)* **11**, 1494-1504, doi:10.1261/rna.7284905 (2005).

43   Chen, G., Wen, J. D. & Tinoco, I., Jr. Single-molecule mechanical unfolding and folding of a pseudoknot in human telomerase RNA. *RNA (New York, N.Y.)* **13**, 2175-2188, doi:10.1261/rna.676707 (2007).

44   Li, P. T., Collin, D., Smith, S. B., Bustamante, C. & Tinoco, I., Jr. Probing the mechanical folding kinetics of TAR RNA by hopping, force-jump, and force-ramp methods. *Biophys J* **90**, 250-260, doi:10.1529/biophysj.105.068049 (2006).

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

46   Leger, M., Sidani, S. & Brakier-Gingras, L. A reassessment of the response of the bacterial ribosome to the frameshift stimulatory signal of the human immunodeficiency virus type 1. *RNA (New York, N.Y.)* **10**, 1225-1235, doi:10.1261/rna.7670704 (2004).

47   Horsfield, J. A., Wilson, D. N., Mannering, S. A., Adamski, F. M. & Tate, W. P. Prokaryotic ribosomes recode the HIV-1 gag-pol-1 frameshift sequence by an E/P site post-translocation simultaneous slippage mechanism. *Nucleic Acids Res* **23**, 1487-1494, doi:10.1093/nar/23.9.1487 (1995).

48   Brilot, A. F., Korostelev, A. A., Ermolenko, D. N. & Grigorieff, N. Structure of the ribosome with elongation factor G trapped in the pretranslocation state. *Proc Natl Acad Sci U S A* **110**, 20994-20999, doi:10.1073/pnas.1311423110 (2013).

49   Lin, J., Gagnon, M. G., Bulkley, D. & Steitz, T. A. Conformational changes of elongation factor G on the ribosome during tRNA translocation. *Cell* **160**, 219-227, doi:10.1016/j.cell.2014.11.049 (2015).

50   Bulkley, D. *et al.* The antibiotics dityromycin and GE82832 bind protein S12 and block EF-G-catalyzed translocation. *Cell Rep* **6**, 357-365, doi:10.1016/j.celrep.2013.12.024 (2014).

51   Fislage, M. *et al.* Cryo-EM shows stages of initial codon selection on the ribosome by aa-tRNA in ternary complex with GTP and the GTPase-deficient EF-TuH84A. *Nucleic Acids Res* **46**, 5861-5874, doi:10.1093/nar/gky346 (2018).

52   Loveland, A. B., Demo, G. & Korostelev, A. A. Cryo-EM of elongating ribosome with EF-Tu*GTP elucidates tRNA proofreading. *Nature*, doi:10.1038/s41586-020-2447-x (2020).

53   Hussain, T., Llacer, J. L., Wimberly, B. T., Kieft, J. S. & Ramakrishnan, V. Large-Scale Movements of IF3 and tRNA during Bacterial Translation Initiation. *Cell* **167**, 133-144 e113, doi:10.1016/j.cell.2016.08.074 (2016).

54   McCauley, M. J., Rouzina, I., Li, J., Nunez, M. E. & Williams, M. C. Significant Differences in RNA Structure Destabilization by HIV-1 GagDp6 and NCp7 Proteins. *Viruses* **12**, doi:10.3390/v12050484 (2020).

55   Liu, T. *et al.* Direct measurement of the mechanical work during translocation by the ribosome. *eLife* **3**, e03406, doi:10.7554/eLife.03406 (2014).

56   Kim, H. K. *et al.* A frameshifting stimulatory stem loop destabilizes the hybrid state and impedes ribosomal translocation. *Proc Natl Acad Sci U S A* **111**, 5538-5543, doi:10.1073/pnas.1403457111 (2014).

57   Qin, P., Yu, D., Zuo, X. & Cornish, P. V. Structured mRNA induces the ribosome into a hyper-rotated state. *EMBO Rep* **15**, 185-190, doi:10.1002/embr.201337762 (2014).

58   Bao, C. *et al.* mRNA stem-loops can pause the ribosome by hindering A-site tRNA binding. *eLife* **9**, doi:10.7554/eLife.55799 (2020).

59   Zhou, J., Lancaster, L., Donohue, J. P. & Noller, H. F. Spontaneous ribosomal translocation of mRNA and tRNAs into a chimeric hybrid state. *Proc Natl Acad Sci U S A* **116**, 7813-7818, doi:10.1073/pnas.1901310116 (2019).

60   Rhoads, R. E. eIF4E: new family members, new binding partners, new roles. *J Biol Chem* **284**, 16711-16715, doi:10.1074/jbc.R900002200 (2009).

61   Pestova, T. V., Hellen, C. U. & Shatsky, I. N. Canonical eukaryotic initiation factors determine initiation of translation by internal ribosomal entry. *Molecular and cellular biology* **16**, 6859-6869, doi:10.1128/mcb.16.12.6859 (1996).

62   Pestova, T. V., Shatsky, I. N. & Hellen, C. U. Functional dissection of eukaryotic initiation factor 4F: the 4A subunit and the central domain of the 4G subunit are sufficient to mediate internal entry of 43S preinitiation complexes. *Molecular and cellular biology* **16**, 6870-6878, doi:10.1128/mcb.16.12.6870 (1996).

63   Neidel, S. *et al.* Vaccinia virus protein A49 is an unexpected member of the B-cell Lymphoma (Bcl)-2 protein family. *J Biol Chem* **290**, 5991-6002, doi:10.1074/jbc.M114.624650 (2015).

64   Winter, G. *xia2*: an expert system for macromolecular crystallography data reduction. *J. Appl. Cryst.* **43**, 186-190 (2009).

65   Kabsch, W. XDS. *Acta Crystallogr D Biol Crystallogr* **66**, 125-132, doi:10.1107/S0907444909047337 (2010).

66   Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr* **69**, 1204-1214, doi:10.1107/S0907444913000061 (2013).

67   Karplus, P. A. & Diederichs, K. Linking crystallographic model and data quality. *Science* **336**, 1030-1033, doi:10.1126/science.1218231 (2012).

68   Vonrhein, C., Blanc, E., Roversi, P. & Bricogne, G. Automated structure solution with autoSHARP. *Methods in molecular biology (Clifton, N.J.)* **364**, 215-230, doi:10.1385/1-59745-266-1:215 (2007).

69   Sheldrick, G. M. A short history of SHELX. *Acta Crystallogr A* **64**, 112-122, doi:10.1107/S0108767307043930 (2008).

70   Abrahams, J. P. & Leslie, A. G. Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallogr D Biol Crystallogr* **52**, 30-42, doi:10.1107/S090744995008754 (1996).

71   Perrakis, A., Harkiolaki, M., Wilson, K. S. & Lamzin, V. S. ARP/wARP and molecular replacement. *Acta Crystallogr D Biol Crystallogr* **57**, 1445-1450, doi:10.1107/s0907444901014007 (2001).

72   McCoy, A. J. *et al.* Phaser crystallographic software. *Journal of applied crystallography* **40**, 658-674, doi:10.1107/S0021889807021206 (2007).

73   Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501, doi:10.1107/S0907444910007493 (2010).

74   Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221, doi:10.1107/S0907444909052925 (2010).

75   Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12-21, doi:10.1107/S0907444909042073 (2010).

76   Dolinsky, T. J., Nielsen, J. E., McCammon, J. A. & Baker, N. A. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res* **32**, W665-667, doi:10.1093/nar/gkh381 (2004).

77   Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A* **98**, 10037-10041, doi:10.1073/pnas.181342398 (2001).

78   Krissinel, E. & Henrick, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr* **60**, 2256-2268, doi:10.1107/S0907444904026460 (2004).

79   Holm, L. & Laakso, L. M. Dali server update. *Nucleic Acids Res* **44**, W351-355, doi:10.1093/nar/gkw357 (2016).

80   Redfern, O. C., Harrison, A., Dallman, T., Pearl, F. M. & Orengo, C. A. CATHEDRAL: a fast and effective algorithm to predict folds and domain boundaries from multidomain protein structures. *PLoS Comput Biol* **3**, e232, doi:10.1371/journal.pcbi.0030232 (2007).

81   Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774-797, doi:10.1016/j.jmb.2007.05.022 (2007).

82   Stephenson, W., Wan, G., Tenenbaum, S. A. & Li, P. T. Nanomanipulation of single RNA molecules by optical tweezers. *J Vis Exp*, doi:10.3791/51542 (2014).

83   Bellaousov, S., Reuter, J. S., Seetin, M. G. & Mathews, D. H. RNAstructure: Web servers for RNA secondary structure prediction and analysis. *Nucleic Acids Res* **41**, W471-474, doi:10.1093/nar/gkt290 (2013).

84   Fixsen, S. M. & Howard, M. T. Processive selenocysteine incorporation during synthesis of eukaryotic selenoproteins. *J Mol Biol* **399**, 385-396, doi:10.1016/j.jmb.2010.04.033 (2010).

85   Powell, M. L., Brown, T. D. & Brierley, I. Translational termination-re-initiation in viral systems. *Biochem Soc Trans* **36**, 717-722, doi:10.1042/BST0360717 (2008).

86   Milon, P. *et al.* Transient kinetics, fluorescence, and FRET in studies of initiation of translation in bacteria. *Methods Enzymol* **430**, 1-30, doi:10.1016/S0076-6879(07)30001-3 (2007).

87   Rodnina, M. V., Semenkov, Y. P. & Wintermeyer, W. Purification of fMet-tRNA(fMet) by fast protein liquid chromatography. *Anal Biochem* **219**, 380-381, doi:10.1006/abio.1994.1282 (1994).

88   Kothe, U., Paleskava, A., Konevega, A. L. & Rodnina, M. V. Single-step purification of specific tRNAs by hydrophobic tagging. *Anal Biochem* **356**, 148-150, doi:10.1016/j.ab.2006.04.038 (2006).

89   Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol* **185**, 60-89, doi:10.1016/0076-6879(90)85008-c (1990).

90   Passmore, L. A. & Russo, C. J. Specimen Preparation for High-Resolution Cryo-EM. *Methods Enzymol* **579**, 51-86, doi:10.1016/bs.mie.2016.04.011 (2016).

91   Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature methods* **14**, 331-332, doi:10.1038/nmeth.4193 (2017).

92   Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J Struct Biol* **192**, 216-221, doi:10.1016/j.jsb.2015.08.008 (2015).

93   Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, doi:10.7554/eLife.42166 (2018).

94   Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of computational chemistry* **25**, 1605-1612, doi:10.1002/jcc.20084 (2004).