Coordination of -1 programmed ribosomal frameshifting by transcript and nascent chain features revealed by deep mutational scanning

Patrick J. Carmody[1,†], Matthew H. Zimmer[2,†], Charles P. Kuntz[1], Haley R. Harrington[1], Kate E. Duckworth[1], Wesley D. Penn[1], Suchetana Mukhopadhyay[3,*], Thomas F. Miller, III[2,*] and Jonathan P. Schlebach [1,*]

[1]Department of Chemistry, Indiana University, Bloomington, IN 47405, USA, [2]Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125, USA and [3]Department of Biology, Indiana University, Bloomington, IN 47405, USA

ABSTRACT

Programmed ribosomal frameshifting (PRF) is a translational recoding mechanism that enables the synthesis of multiple polypeptides from a single transcript. During translation of the alphavirus structural polyprotein, the efficiency of −1PRF is coordinated by a 'slippery' sequence in the transcript, an adjacent RNA stem–loop, and a conformational transition in the nascent polypeptide chain. To characterize each of these effectors, we measured the effects of 4530 mutations on −1PRF by deep mutational scanning. While most mutations within the slip-site and stem–loop reduce the efficiency of −1PRF, the effects of mutations upstream of the slip-site are far more variable. We identify several regions where modifications of the amino acid sequence of the nascent polypeptide impact the efficiency of −1PRF. Molecular dynamics simulations of polyprotein biogenesis suggest the effects of these mutations primarily arise from their impacts on the mechanical forces that are generated by the translocon-mediated cotranslational folding of the nascent polypeptide chain. Finally, we provide evidence suggesting that the coupling between cotranslational folding and −1PRF depends on the translation kinetics upstream of the slip-site. These findings demonstrate how −1PRF is coordinated by features within both the transcript and nascent chain.

GRAPHICAL ABSTRACT
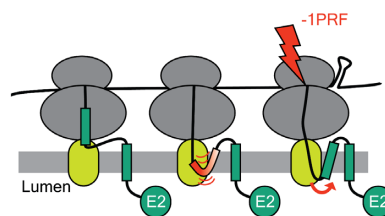
INTRODUCTION

Programmed ribosomal frameshifting (PRF) is a translational recoding mechanism that occurs in all kingdoms of life. Though a handful of prokaryotic and eukaryotic PRF motifs have been identified, most of the well-characterized motifs are found within viral genomes (1). Viruses utilize ribosomal frameshifting to increase their genomic coding capacity and to regulate the stoichiometric ratios of viral protein synthesis. Some rely on these motifs to coordinate genomic replication, while others utilize PRF to regulate the production of the structural proteins that mediate assembly (2). In some cases, the frameshift products are themselves virulence factors that antagonize the host interferon response (3–6). For these reasons, the efficiency of PRF, which is globally regulated by both host and viral proteins (7,8), is often critical for infection and immunity.

The most common type of PRF involves a -1 shift in reading frame (−1PRF) and minimally requires a 'slippery' heptanucleotide site in the transcript that typically takes the form $X_1 XXY_4 YYZ_7$, where X represents three identical nucleotides, Y is an A or U, and Z can be an A, C or U (9). This structure reduces the thermodynamic penalty as-

sociated with non-native codon-anticodon interactions in the -1 reading frame (10), which increases the intrinsic efficiency of ribosomal frameshifting by several orders of magnitude (11,12). Most −1PRF motifs also feature an RNA stem–loop or pseudoknot that forms 6–8 bases downstream of the slip-site (13,14). This structure imposes a mechanical resistance that pauses the ribosome on the slip-site and provides time for the anticodon loops of the A- and/ or P-site tRNA to sample alternative base pairing interactions (15–18). Frameshifting is also enhanced by the depletion of the tRNA pool that decodes the YYZ_7 codon ('hungry' frameshifting), which provides the opportunity for the P-site tRNA to move out of frame (19). In addition to mRNA and tRNA effectors, there is growing evidence to suggest that −1PRF can be tuned by the conformational properties of the nascent polypeptide chain (20–23). Our group recently found that −1PRF is enhanced by mechanical forces generated by the cotranslational folding of the nascent polypeptide chain (22). This finding has since been echoed by recent evidence suggesting cotranslational folding of nsp10 alters −1PRF during translation of SARS-CoV-2 genome (23). The efficiency of −1PRF can also be modified in *trans* by a variety of host and viral proteins and nucleic acids (2). Thus, the efficiency of ribosomal frameshifting can be modulated by a wide variety of effectors that tune the rate of translation, the structural properties of the transcript, and/ or the conformation of the nascent chain. Nevertheless, a unified understanding of how these features cooperate to establish the efficiency of PRF is needed (24).

In this work, we utilize deep mutational scanning (DMS) to identify sequence constraints associated with −1PRF in the multipass integral membrane protein known as the Sindbis virus (SINV) structural polyprotein. We report the effects of 4530 mutations on the relative efficiency of −1PRF in the context of an extended genetic reporter containing the native features that coordinate the biosynthesis and cotranslational folding of the nascent polyprotein at the endoplasmic reticulum (ER) membrane (Figure 1A). Incorporating these elements into our reporter facilitated the identification of −1PRF modulators in both the transcript and within several regions of the nascent chain. As expected, mutational patterns within the slip-site and the region immediately downstream of the slip-site are consistent with the formation of a previously characterized RNA stem–loop (25,26). However, we find that −1PRF is also sensitive to numerous mutations within the ~160 base region preceding the slip-site, which encode a pair of transmembrane (TM) domains that occupy the ribosomal exit tunnel and translocon during frameshifting. In conjunction with a suite of atomistic and coarse-grained molecular dynamics (CGMD) simulations (27), we provide evidence suggesting that most mutations within this region attenuate −1PRF through their effects on a folding intermediate that forms during translocon-mediated membrane integration of the nascent polyprotein. We also find that −1PRF appears to be sensitive to changes in the translation kinetics within the region immediately preceding the slip-site, which influences the pulling forces generated by the cotranslational folding of the nascent chain. Together, these findings provide a detailed description of the sequence con-

straints of −1PRF within the alphavirus structural polyprotein.

MATERIALS AND METHODS

Plasmid and library preparation

−1PRF measurements were made using a previously described fluorescent reporter that selectively generates a fluorescent mKate protein as a result of −1PRF during translation of the SINV 6K gene (22). To control for variations in expression, the intensity of the mKate −1PRF signal is measured in relation to the eGFP intensity generated by a downstream IRES-eGFP cassette. To adapt this reporter for DMS, the CMV promoter and enhancer within this vector were replaced with an attB recombination site using NEBuilder. An Nt.BbvCI/ Nb.BbvCI endonuclease site was also introduced into the backbone using site-directed mutagenesis in order to facilitate the production of one-pot genetic libraries. A library of single-codon variants was generated using nicking mutagenesis (28). Briefly, the sense strand of the parental plasmid was nicked and degraded. The antisense strand was then used as template for the synthesis of a complementary strand using a pool of mutagenic oligonucleotides that each contain a randomized NNN codon. The WT antisense strand of the heteroduplex was nicked and degraded. The remaining mutagenized sense strand library was then used as a template for synthesis using a universal primer that anneals outside of the mutagenized region. The resulting plasmid library was then electroporated into electrocompetent NEB10β cells (New England Biolabs, Waltham, MA), which were then grown in liquid culture prior to purification of the plasmid library using a ZymoPure II Endotoxin-Free Midiprep kit (Zymo Research, Irvine, CA). An analysis of our final DMS data suggests this library contains at least 7,614 of 7,686 possible variants. Nevertheless, in this manuscript we have chosen to focus on the effects of 4530 of these mutations within specific regions of interest, which are shown in the heatmaps in Figures 2A, 3 and 6. To validate DMS measurements, a representative series of individual variants were introduced into the originally described expression vector using site-directed mutagenesis (22). To compare −1PRF levels in the context of a dual luciferase reporter, the fragment of the SINV structural polyprotein shown in Figure 1A (residues 265−817) was inserted into a dual-luciferase reporter construct (pJD2257) (29) that was modified in order to remove a cryptic splice site in the Renilla luciferase gene (30).

Cell culture and production of recombinant stable cell lines

All cells were grown in Dulbecco's modified eagle medium (Gibco, Carlsbad, CA) supplemented with 10% fetal calf serum (Corning, Corning, NY), and penicillin (100 U/ml)/ streptomycin (100 μg/ml) at 37°C in a humidified incubator containing 5% CO_2 by volume. Recombinant stable cell libraries expressing a mix of polyprotein reporter variants from a uniform genomic locus were generated by cotransfecting a genetically-modified HEK293T cell line with the plasmid library and a Bxb1 expression vector as was described previously (31,32). Cells that underwent the desired recombination reaction were isolated using a FACS Aria

II cell sorter (BD Biosciences, Franklin Lakes, NJ) based on the characteristic gain in eGFP and loss of blue fluorescent protein (BFP) expression that occurs as a result of the insertion of a single copy of the reporter downstream from the Tet-inducible promoter within the genomic 'landing pad' (31). Following isolation, recombinant cells were grown in media containing 2 μg/ml doxycycline in order to maintain the stable expression of polyprotein variants from the genomic landing pad.

Deep mutational scanning

A minimum of 2 million stably recombined cells were isolated by FACS to ensure the library of stable cells represents an exhaustive sampling of the variants within the plasmid library (∼1000 × coverage). This population was then expanded prior to its fractionation into quartiles (∼2 million cells per fraction) according to the mKate: eGFP intensity ratio by FACS. Each cellular fraction was then expanded and harvested prior to the extraction of genomic DNA using the GenElute Mammalian Genomic DNA Miniprep kit (Sigma-Aldrich, St. Louis, MO). A semi-nested PCR approach was then used to selectively amplify the mutagenized region of the polyprotein reporter from the genomic integration site in a similar manner as was previously described (31,32). Amplicons containing the mutagenized regions were prepared from each cellular isolate and sequenced using paired end reads on a 600 cycle Illumina MiSeq cell (Illumina Inc., San Diego, CA) to a final depth of ∼2 million reads per amplicon/isolate, which were filtered based on quality score. The mKate: eGFP intensity ratio for each variant was then inferred from the number of sequence-based identifications within each cellular isolate and normalized relative to the value of WT, as was previously described (32). It should be noted that, due to the nature of the scoring system, the lowest possible score for mKate: eGFP intensity ratios is 0.35, which corresponds to the average intensity ratio of the lowest FACS bin divided by the intensity ratio of WT from each experiment (32).

Relative mKate: eGFP intensity ratios for codon substitutions were mapped according to the relative abundance of decoding tRNAs (Figure 6) based on previous measurements of the relative abundance of tRNAs in HEK293 cells (33). Briefly, we averaged previously reported replicate measurements of the relative abundance of tRNAs in HEK293 cells. These values were then used to calculate a score based on the sum of the abundances for all tRNAs that are capable of decoding each codon.

−1PRF reporter measurements

In order to validate DMS measurements, we characterized a series of individual variants using both a fluorescence-based reporter (22) and a dual-luciferase reporter for −1PRF in the SINV 6K gene. Fluorescence reporters bearing polyprotein variants of interest were transiently expressed in HEK293T cells, and the relative mKate: eGFP intensity ratios were determined by flow cytometry as previously described (22). HEK293T cells were transiently transfected with dual-luciferase reporter constructs using lipofectamine 3000 (Invitrogen, Carslbad, CA) in accordance

with the manufacturer's instructions. Cells were harvested two days post-transfection and lysed using the Passive Lysis Buffer provided in the Dual-Luciferase Reporter Assay System (Promega, Madison, WI). This lysis buffer contains CHAPS detergent that should solubilize any remaining uncleaved, membrane bound reporter enzymes. The relative activities of the firefly luciferase and renilla luciferase domains in the clarified lysate were then measured using the Dual-Luciferase Reporter Assay System on a Synergy Neo2 plate reader (Biotek, Winooski, VT) in accordance with the manufacturer's instructions. To determine the −1PRF efficiency, the firefly: renilla luciferase ratio was normalized relative to that of a construct lacking the insert as was previously described (34).

Computational predictions of topological energetics

The effects of mutations on the free energy associated with the transfer of the second TM domain of the E2 protein (TM2) from the translocon to the ER membrane were estimated using the ΔG predictor (35) using a series of individual predictions for TM2 variants assuming a WT sequence of LTPYALAPNAVIPTSLALLCCVR.

Coarse-grained simulations of polyprotein translation

Coarse-grained simulations use a previously developed and validated methodology (22,36,37). Simulations are based on coarsened representations of the ribosome exit tunnel, Sec translocon, and nascent chain. The nascent chain is represented as a polymer of beads, each of which represents three amino acids. Each bead has a hydrophobicity and charge derived from its three constituent amino acids. The solvent and lipid bilayer are modelled implicitly. The lateral gate of the Sec translocon stochastically switches between open and closed conformations with probability dependent on the free energy difference between the two conformations. The structure of the ribosome and Sec translocon are based on cryo-EM structures (38), and aside from the opening/closing of the lateral gate, are fixed in place during simulation.

Model parameterization is unchanged from previously published work (22,36). Beads are added to the nascent chain at a rate of 5 amino acids per second unless otherwise specified. Integration is carried out using overdamped Langevin dynamics with a timestep of 300 ns and a diffusion coefficient of 253 nm^2/s. Translation starts 33 amino acids before the N-terminus of TM1 of the E2 protein (P663).

To measure pulling forces, translation is halted when the ribosome occupies the slip site. The dynamics of the nascent chain are then allowed to evolve for three seconds, during which time the forces exerted on the N-terminal bead are measured every 3 ms. Because the exit tunnel is truncated in the CGMD model, we add 27 amino acids to the index of the final bead to account for the unmodeled nascent chain in the omitted part of the exit tunnel. All single mutants from positions 725 to 754 are simulated. Each mutant was simulated in 100 independent runs in order to ensure the estimate for the mean pulling force has low statistical error. Data from trajectories in which TM1 does not adopt the correct N_{out} topology were not included in the analysis.

The ratios of topomeric isomers were obtained independently from the force-measuring simulations. To observe the final topology of TM2 in the membrane, the polyprotein was translated from amino acid P663 to I837 without pausing. Restraints were placed on TM1 to ensure it adopts the correct topology. No restraints were placed on TM2. Upon translation of the final bead, the nascent chain was released from the ribosome and simulation continued. After 3 s, the topology of the protein was recorded. Simulations were repeated 100 times to obtain the probability of TM2 integrating per mutant.

Atomistic simulations of polyprotein TM2 in the translocon

An atomistic model was built of polyprotein TM2 in the Sec translocon using the cryo-EM structure with PDB code of 6ITC as an initial template (39). This structure captures a polypeptide entering a lipid bilayer through the lateral gate of SecY. The translocating peptide sequence in the structure was replaced with the polyprotein sequence using MODELLER (40), aligning the TM2 helix in the polyprotein with the helical pro-OmpA signal sequence that is exposed to the lipid bilayer in the cryoEM structure. Specifically, residue M1 of the peptide was replaced with residue E723 of the polyprotein. Two alternative structures were generated by shifting the mapping of TM2 amino acids two amino acids forwards. Specifically, residue M1 of the peptide was replaced with either residues L725 or P727. Unresolved amino acids in the translocating peptide were modeled in, also using MODELLER. The antibodies and GFP present in the cryoEM structure were removed. The resulting structure was then solvated with TIP3P water and POPC lipids, filling a simulation box that extended 10 Å beyond the protein complex. The structure was energetically locally minimized, and then the α carbons in SecA more than 15 Å from the translocating peptide were fixed in place with 1 kcal/mol/Å^2 harmonic restraints. Molecular dynamics simulations were then run for 150 ns with a 1.5 fs timestep at 300 K and 1 atm using Desmond (41). Root mean square fluctuations of the amino acid alpha carbons were calculated using the final 100 ns of the trajectory.

RESULTS

Production of a stable HEK293T cell line expressing a programmed array of −1PRF reporter variants

To probe the sequence constraints of −1PRF, we employed DMS to map the effects of mutations on ribosomal frameshifting within the SINV structural polyprotein. Mutagenic effects were assessed in the context of a reporter that selectively generates a fluorescent protein (mKate) as a result of −1PRF during translation and processing of the SINV structural polyprotein (Figure 1A). The design of this reporter is similar, in principle, to those described in a recent high-throughput investigation of PRF (24), except that we included a much larger fragment of the transcript (1.7 kb, 557 amino acids) and opted to normalize for changes in expression using a 3′ IRES-eGFP cassette rather than with a 5′ fusion domain. Appending a 5′ luciferase-P2A cassette to the E3 gene appears to compromise the
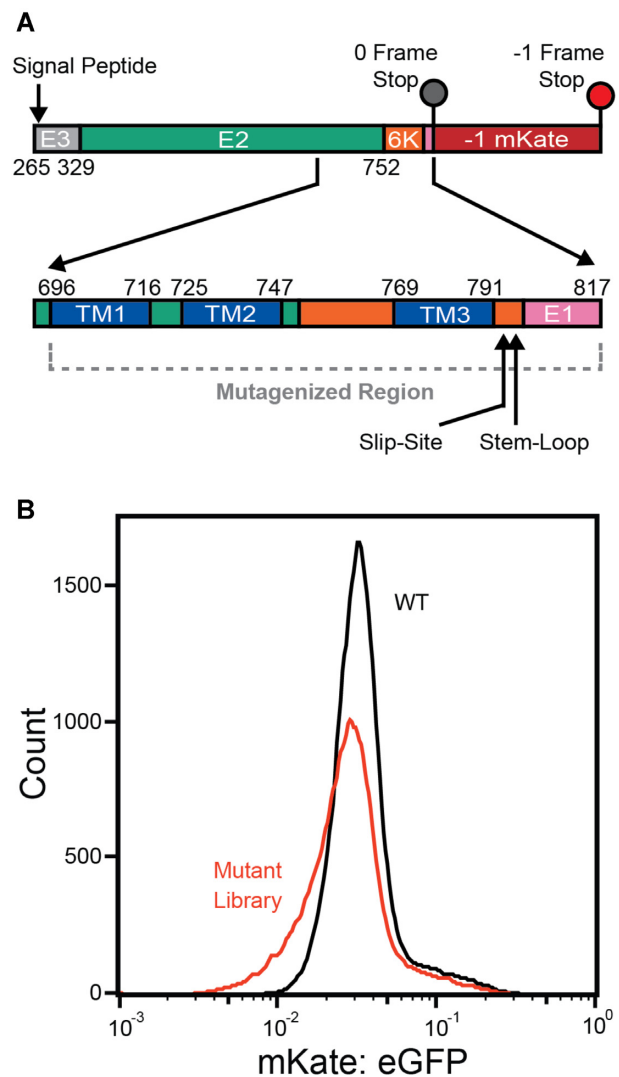

Figure 1. Library of ribosomal frameshifting reporter variants. (**A**) A cartoon depicts the structure of a reporter that generates a fluorescent protein (mKate) as a result of −1PRF in the SINV structural polyprotein (top). A library of reporter variants was generated by randomizing every 0-frame codon within a region ranging from residues 696−817 (bottom). (**B**) A histogram depicts the range of single-cell mKate: GFP intensity ratios among recombinant stable cells expressing the WT reporter (black) or the library of reporter variants (red).

translocon-mediated membrane integration of TM2 (Supplementary Figure S1)—a key conformational transition that stimulates −1PRF (22). Thus, this design preserves the integrity of the native signal peptide in the E3 protein (Figure 1A), which is critical for the targeting and maturation of the nascent polyprotein at the ER membrane. To survey the effects of mutations on the reporter signal, we first generated a one-pot genetic library by randomizing each 0-frame codon within a region beginning at the N-terminal codon of the first TM domain (TM1) of the E2 protein (I696) and ending 23 codons downstream of the slip-site (P817, Figure 1A). We then used this genetic library (7614 total variants) to generate recombinant stable HEK293T cell lines in which each cell expresses a single variant from a unique Tet-inducible promoter within the genome, as previously
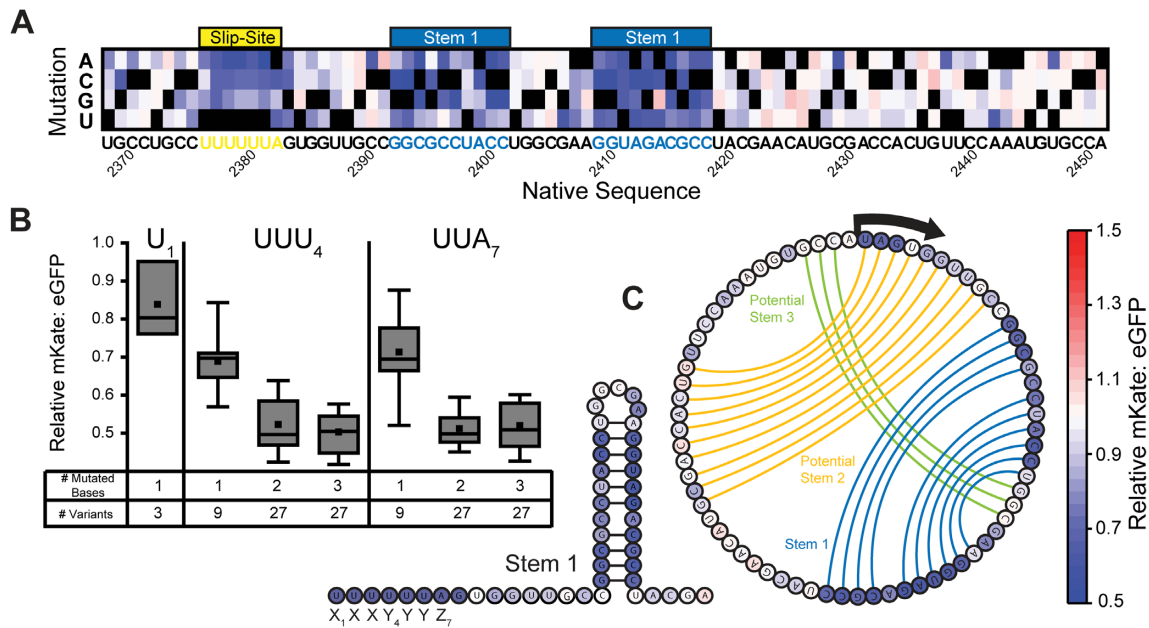
Figure 2. Impact of slip-site and stem–loop mutations on ribosomal frameshifting. (**A**) A heatmap depicts the effects of every single nucleotide substitution (Y-coordinate) at each position (X-coordinate) within or near the slip-site on the mKate: eGFP ratio as determined by DMS. Ratios were normalized relative to that of WT, and black squares are shown in place of the native nucleobase at each position. The relative position of the slip-site (yellow) and stem 1 (blue) are shown for reference. (**B**) A box plot depicts the distribution of relative mKate: eGFP values associated with variants bearing one, two, or three mutations at certain positions within the slip-site. The upper and lower boundaries of the box correspond to the 75th and 25th percentile values, and the upper and lower whiskers correspond to the 90th and 10th percentile values, respectively. The horizontal line within the box corresponds to the median value, and the square corresponds to the position of the average value. (**C**) A secondary structure diagram shows the predicted structure of stem 1, and a circular contact plot depicts previously predicted base pairing interactions within stem 1 (blue lines), potential stem 2 (orange lines), and potential stem 3 (green lines) (25). Nucleotides are colored according to the average relative mKate: eGFP intensity ratio across the three mutations at each position.

described (31). Importantly, the specificity of this recombination reaction ensures individual variants are expressed at similar levels within each cell. Nevertheless, we normalized mKate levels according to the IRES-eGFP intensity to account for small variations in expression, as has been previously described (42).

On average, stable cells expressing individual single-codon variants have an mKate: eGFP intensity ratio that is comparable to that of cells expressing the WT reporter (Figure 1B). However, a sub-set of these cells express variants that generate intensity ratios outside the WT intensity range (Figure 1B). This suggests that, while most mutations have minimal impact, the library contains many that modify the efficiency of −1PRF. To analyze the effects of individual mutations, we fractionated these cells according to their mKate: eGFP intensity ratios using fluorescence activated cell sorting (FACS), quantified the variants within each fraction by deep sequencing, then used these data to estimate the intensity ratio for each individual variant as described previously (32). Intensity ratios were generally consistent across biological replicates when normalized according to the WT value (Pearson's $R = 0.79−0.83$, Supplementary Figure S2A−C). Intensity ratios determined by DMS were also correlated with independent measurements in cells transiently expressing individual reporter variants (Supplementary Figure S2D). In the following, we provide a focused analysis of the effects of 4530 mutations that include a comprehensive set of missense variants upstream of the slip-site, codon-level substitutions within the exit tunnel

region, single, double, and triple mutants within the slip-site, and a collection of single nucleotide variants in the downstream region encoding the stimulatory RNA structure. This subset of variants collectively reveals how structural features within the nascent chain and RNA coordinate the efficiency of ribosomal frameshifting.

Sequence constraints within the heptanucleotide slip-site

Slippery nucleotide sequences enable favorable codon-anticodon base pairing in the -1 reading frame, and the efficiency of −1PRF is therefore highly sensitive to mutations that disrupt the slip-site (43). As expected, we find that every single-nucleotide variant (SNV) within the heptanucleotide slip-site (U_1 UUU_4 UUA_7) measurably decreases −1PRF (Figure 2A, yellow). Furthermore, an analysis of 128 single, double, and triple mutants within the slip-site reveals that frameshifting decreases with an increasing mutational load within the UUU_4 (P-site) or UUA_7 (A-site) codons (Figure 2B). These mutations presumably decrease the efficiency of −1PRF by creating unfavorable base pairing between the transcript and anticodon loop in the −1 frame. However, we also find frameshifting to be sensitive to mutations within the adjacent codon upstream of the slip-site (Figure 2A). This observation potentially suggests that frameshifting is sensitive to structural perturbations within the ribosomal E-site, or that the observed PRF signal may arise from some combination of −1 and −4 frameshifting transitions as has been observed for other PRF motifs (44). It is also possible
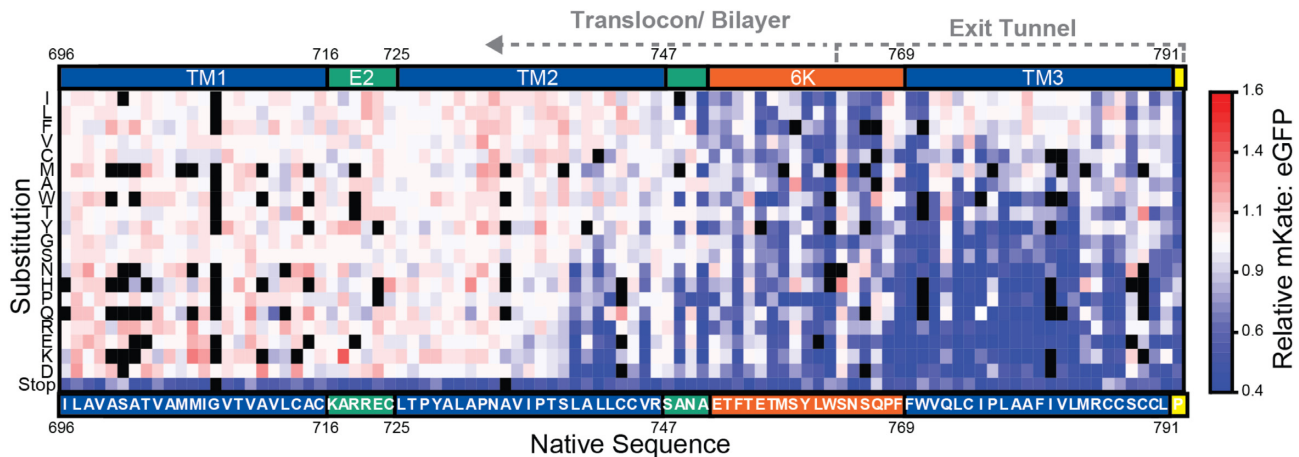
Figure 3. Impact of nascent chain mutations on ribosomal frameshifting. A heatmap depicts the effects of every amino acid substitution (Y-coordinate) at each position (X-coordinate) within the nascent polypeptide chain on the mKate: eGFP ratio as determined by DMS. Substitutions are arranged from most hydrophobic (top) to most polar (bottom). Ratios were normalized relative to that of WT, and black squares indicate mutations that lack coverage. The positions of each TM domain (blue) and the slip-site (yellow) as well as the domain boundaries for the E2 (green) and 6K (orange) proteins are shown for reference. The amino acid that corresponds to the first codon of the slip site (containing U1) is indicated in yellow. The residues that are likely to reside within the exit tunnel or outside of the ribosome during ribosomal frameshifting are indicated with dashed lines.

that the observed −1PRF signal arises from a mix of canonical frameshifting (two-tRNA) and 'hungry' frameshifting (one-tRNA), where the P-site tRNA frameshifts prior to the decoding of the UUA_7 codon. The overall efficiency of both the HIV and Semliki forest virus −1PRF motifs is maximized when the relative abundance of the tRNA that decodes the UUA_7 codon is depleted (19). Similarly, we find that −1PRF is decreased by all three A_7 mutations in the SINV slip-site (Figure 2A), which should attenuate hungry frameshifting given that these codons are decoded by tRNA that are more abundant in these cells (33,45). Together, these observations highlight the essential role of the slip-site and its adjacent bases, yet suggest the observed −1PRF signal potentially arises from a spectrum of frameshifting transitions.

Sequence constraints within the stimulatory RNA stem–loop

Efficient ribosomal frameshifting requires a stimulatory RNA secondary structure that impedes translocation and increases the dwell time of the ribosome on the slip-site. Stimulatory structures within the sub-genomic RNA of a number of alphaviruses have been previously characterized (25,26,29). Secondary structure predictions for SINV suggest the region adjacent to the slip-site is capable of forming a pseudoknot consisting of one primary stem–loop (stem 1), a secondary stem–loop within the spacer region (potential stem 2), and a short stem that basepairs within the loop of stem 1 (potential stem 3) (25). However, experimental evidence suggests frameshifting is primarily driven by stem 1 (25). Consistent with these findings, our map of mutational effects shows that frameshifting is highly sensitive to mutations within a region 8 bases downstream of the slip-site, beginning with a pair of C→A mutations at bases 2390 and 2391 (Figure 2A). This region generally corresponds to the position of stem 1 (Figure 2C). Furthermore, the contacts suggested by this map correspond exactly to those identified by SHAPE analysis (26). Interestingly, stem 1 contains

a single mismatch (Figure 2C), and PRF is enhanced by the only SNV that creates a G-C base pair at this position (2414 A→G) but not by the mutation that creates a weaker A-U base pair (2396 C→U, Figure 2A). This observation implies that frameshifting is sensitive to changes in the overall stability of stem 1. We also find that frameshifting is enhanced by two SNVs that extend the base of the stem–loop by an additional base pair (2391 C→A, 2419 U→ G, Figure 2A). Together, these observations confirm that stem 1 is critical for PRF efficiency. While more stable equilibrium structure(s) could potentially form, our functional analysis suggests stem 1 is the dominant structure within the ensemble that stimulates −1PRF between rounds of ribosome-mediated unwinding (46,47).

Sequence constraints within the nascent polypeptide chain

Our recent investigations of the impact of the nascent chain on −1PRF revealed that the mechanical force generated by the translocon-mediated membrane integration of the second TM domain within the E2 protein enhances ribosomal frameshifting in the SINV structural polyprotein (22). Our DMS results reveal that ribosomal frameshifting is sensitive to a variety of mutations that alter the amino acid sequence within the portion of the nascent chain that has been translated at the point of frameshifting, including mutations to residues within the exit tunnel, between the exit tunnel and translocon, and within TM2 of the E2 protein (Figure 3). Synonymous mutations upstream of the region encoding exit tunnel residues have minimal impact on ribosomal frameshifting relative to the effects of missense mutations (Supplementary Figure S3). Furthermore, mutagenic effects upstream of the slip-site are lost upon deletion of the TM domains in E2 (Supplementary Figure S4), which suggests that the observed deviations in −1PRF arise from the impact of these mutations on translocon-mediated cotranslational folding. Consistent with previous findings (22), our results show that the introduction of polar
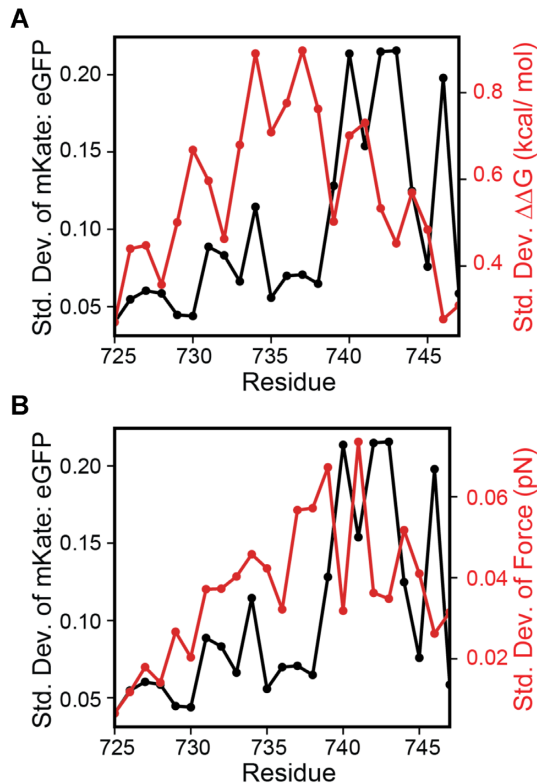
Figure 4. Impact of TM2 mutations on −1PRF, membrane integration energetics, and pulling forces. The standard deviation of the relative mKate: eGFP intensity ratios for all 19 amino acids substitutions (black) at each residue are plotted in relation to the corresponding standard deviation of (A) the predicted change in the transfer free energy associated with the translocon-mediated membrane integration of TM2 as determined by the ΔG predictor (red) (35) and B) the relative pulling force generated by its translocon-mediated membrane integration as was determined by CGMD simulations (red).

or charged side chains within a certain region of TM2 reduces −1PRF (V735-V746) while mutations that enhance the hydrophobicity of TM2 generally increase −1PRF (Figure 3). However, the effects of mutations on −1PRF appear to deviate from their predicted effects on the energetics of translocon-mediated membrane integration (Supplementary Figure S5). Mutations to residues near the center of TM2 are predicted to elicit the largest changes in the energetics of translocon-mediated membrane integration (A734−T738, Figure 4A, red), which reflects the depth-dependence of amino acid transfer free energies (35,48). In contrast, variations in mKate: eGFP intensity ratios are most pronounced among variants bearing mutations within the C-terminal residues of TM2 (Figure 4A, black). This discrepancy suggests that, in addition to the hydrophobicity of TM2 and its corresponding transfer free energy, stimulatory pulling forces may hinge upon structural and/ or dynamic constraints associated with its translocon-mediated membrane integration.

To explore the effects of mutations on nascent chain pulling forces, we performed CGMD simulations of polyprotein translation that include coarse-grained representations of the ribosomal exit tunnel and the Sec translocon in an implicit lipid bilayer (36). The nascent polypep-

tide is treated as a polymer of beads (3 AA/ bead) that emerge from the ribosome exit tunnel at the physiological rate of translation (5 AA/s). The hydrophobicity and charge of each bead is derived from its three constituent amino acids. To quantify stimulatory pulling forces, we paused translation for three seconds once the ribosome reaches the slip-site and measured the force on the nascent chain as it explores the environment in, on, and around the translocon as has been previously described (37). Consistent with previous findings (22), simulations of the 437 missense variants bearing mutations within TM2 reveal that mutations that attenuate pulling forces on the nascent chain generally reduce frameshifting; pulling forces are correlated with experimentally derived relative mKate: eGFP intensity ratios (Pearson's $R = 0.48$, Supplementary Figure S6A). Hydrophobic substitutions in TM2 increase the propensity of TM2 to enter the translocon and partition into the membrane in these simulations, which generates higher pulling forces (Supplementary Figure S6B) and enhanced frameshifting (Supplementary Figure S6C). Conversely, polar mutations decrease translocon occupancy, membrane integration of TM2, and pulling forces. Importantly, both pulling forces from CGMD and the observed mKate: eGFP intensity ratios are most sensitive to mutations within the C-terminal residues of TM2 (Figure 4B, red). This agreement suggests stimulatory forces originate from conformational transitions that are mediated by dynamic interactions between TM2, the translocon, and the lipid bilayer.

To evaluate the structural context of mutations that influence −1PRF, we constructed an atomistic model of TM2 nested within the translocon. An initial model was generated by mapping the polyprotein sequence onto a cryo-EM structure of a nascent chain intermediate that adopts an N_{in} topology comparable to that of TM2 in the context of the translocon (39). After a 150 ns equilibration of the translocon-nascent chain complex within an explicit lipid bilayer, TM2 adopts a tilted orientation in which the N-terminal portion of TM2 projects into the lipids while its C-terminal residues remain wedged within the lateral gate of the translocon (Figure 5A). The N-terminal residues that interact with the bilayer are generally more dynamic than the C-terminal residues that interact with the translocon (Figure 5B). TM2 achieves a similar topological orientation and exhibits similar dynamic fluctuations regardless of how the amino acid sequence is initially mapped onto the structural template of the nascent chain (Supplementary Figure S7). While the magnitude of the dynamic fluctuations observed on the time scales of the atomistic simulations is modest relative to coarse-grained simulations, the conformational dynamics observed within both models indicate that fluctuations within the C-terminal residues of TM2 are suppressed by interactions with the translocon (Figure 5C), which is the final portion of the helix to partition into the bilayer. Mutations that enhance the polarity of these residues generally decrease −1PRF (Figure 5A), presumably by strengthening interactions with the translocon and/ or disfavoring the transfer of these residues into the bilayer. Taken together, these observations suggest mutations in TM2 influence pulling forces, and by extension −1PRF, by modifying interactions that form between the translocon and nascent chain.
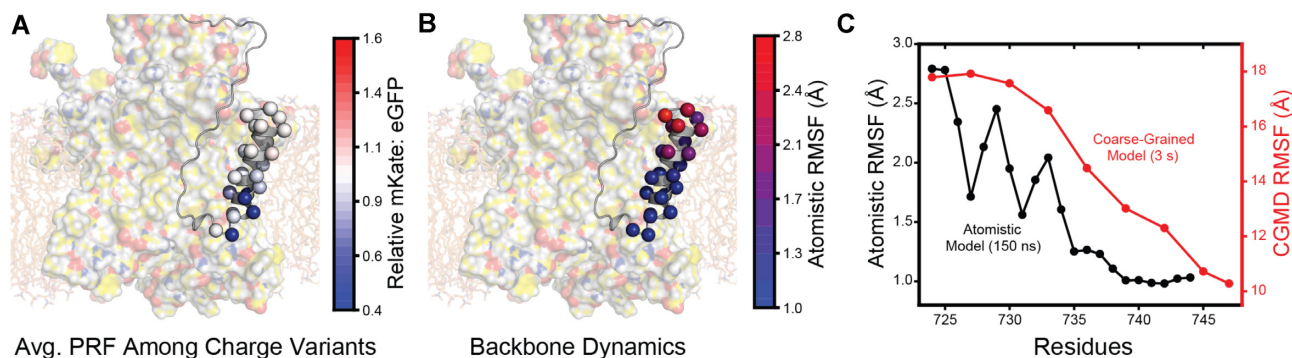
Figure 5. Structural context and conformational dynamics of nascent TM2. (**A**) The average relative mKate: eGFP intensity ratios associated with mutations that introduce charged side chains at each position within TM2 are mapped on to an atomistic model of nascent TM2 at the approximate point of frameshifting. Residues are rendered as spheres and colored based on the relative mKate: eGFP ratio for charged substitutions as was determined by DMS. The image of the nascent chain structure is superimposed on top of the structure of the translocon and lipid bilayer for clarity. (**B**) The root mean square fluctuation (RMSF) of each Cα during a 150 ns all-atom molecular dynamics simulation are mapped on to an atomistic model of nascent TM2 at the approximate point of frameshifting. Residues are rendered as spheres and colored according to RMSF. The image of the nascent chain structure is superimposed on top of the structure of the translocon and lipid bilayer for clarity. (**C**) Fluctuation measurements from atomistic (Cα, black) and coarse-grained (bead indexed to sequence, red) molecular dynamics simulations are plotted for each position within TM2.

Impact of translation kinetics on pulling forces and −1PRF

Ribosomal frameshifting is also highly sensitive to mutations within the region of the transcript encoding the residues between TM2 and the peptidyl transferase center (S748–L791, Figure 3). Most of these mutations alter the portion of the nascent chain that resides within the ribosomal exit tunnel during −1PRF. With respect to the amino acid sequence, these mutagenic trends suggest native −1PRF levels are only maintained by mutations that preserve the hydrophobicity of this segment (Figure 3). This may suggest this segment is capable of forming structural contacts within the exit tunnel (see Discussion). Nevertheless, it should also be noted that the translation of this segment occurs as TM2 begins to emerge from the exit tunnel and interact with the translocon and/or membrane. Modifications that alter the translation kinetics within this region could therefore also impact the membrane integration of TM2 and pulling forces on the nascent chain (27). To evaluate the potential role of translational kinetics, we calculated mKate: eGFP intensity ratios associated with each individual codon substitution, then sorted the values for each codon according to the relative abundance of their decoding tRNAs in HEK293 cells (see Materials and Methods). A heat map of intensity ratios reveals that −1PRF is generally attenuated by mutations within the 21 codons upstream of the slip-site that increase the relative abundance of decoding tRNA (Figure 6). Codons that are decoded by low-abundance tRNAs are generally tolerated within this region (Figure 6), which suggests that efficient −1PRF hinges upon the slow translation of these codons.

Mutations that increase the relative abundance of decoding tRNAs and accelerate translation could potentially decrease −1PRF by reducing the efficiency of translocon-mediated membrane integration (27). To assess how translational kinetics impact the membrane integration of TM2, we carried out a series of CGMD simulations of polyprotein biosynthesis in which the rate of translation was varied. An analysis of these trajectories reveals that the rate of translation has minimal impact on the final topology

of TM2 (Supplementary Figure S8). Nevertheless, −1PRF is likely to be more closely associated with the magnitude of the pulling forces on the nascent chain rather than the equilibrium structure of TM2. To evaluate whether pulling forces are sensitive to translation kinetics, we varied the rate of translation within CGMD simulations and measured the pulling force on the nascent chain while the ribosome occupies the slip-site. One hundred independent trajectories were carried out at distinct translation rates ranging from 1 to 15 amino acids per second, and pulling forces were averaged over 0.8 s once the ribosome reaches the slip-site. A plot of the mean force against translation rate reveals that the pulling force trends upward as translation slows. This observation implies that, when translation is accelerated, the conformational transitions that are required for the generation of stimulatory forces may be slower than the extension of the nascent chain (49). These results provide secondary evidence suggesting the mechanical forces that stimulate −1PRF are likely attenuated at higher translation rates. In conjunction with the mutagenic trends (Figure 6), our findings suggest that frameshifting may be more efficient when translation of the region preceding the slip-site is slow.

DISCUSSION

The interactions between the transcript and translation machinery that stimulate −1PRF have been studied in considerable detail (1,19). More recently, our investigations of the alphavirus structural polyprotein revealed that −1PRF is also sensitive to conformational transitions in the nascent chain (22). The net efficiency of −1PRF in this system is therefore modulated by the interplay between structural features within both the transcript and nascent chain. To map these effectors, we measured the effects of 4530 mutations on the efficiency of −1PRF in the context of a reporter containing a large fragment of the SINV structural polyprotein. This reporter retains all the native features that direct the targeting, processing, and cotranslational folding of the nascent chain (including the signal peptide and
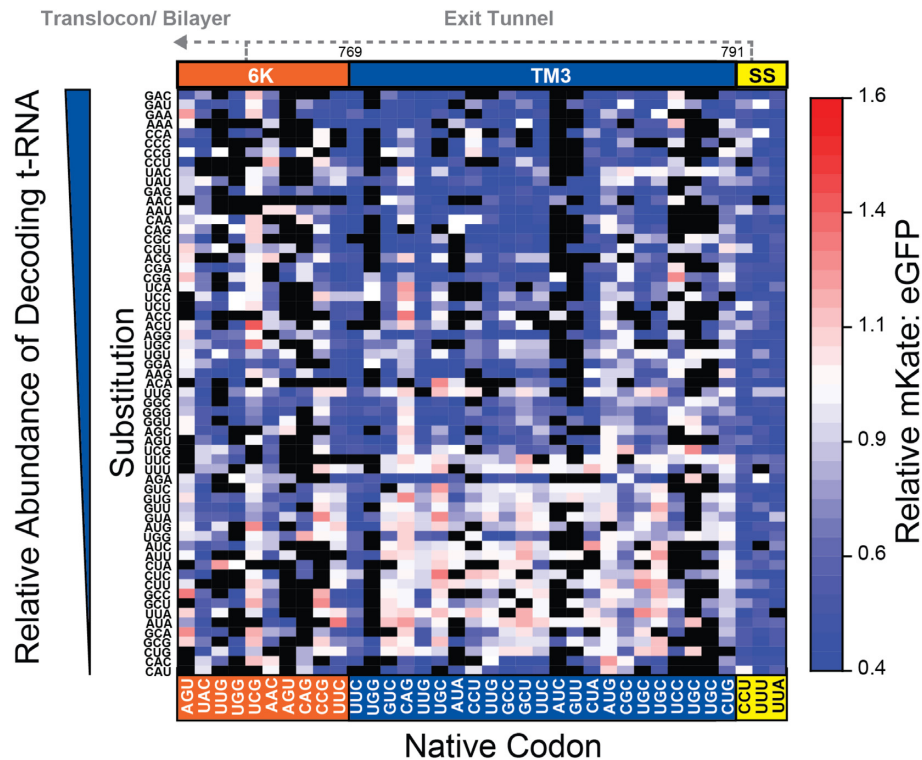
Figure 6. Impact of codon modifications on ribosomal frameshifting. A heatmap depicts the effects of every codon substitution (Y-coordinate) at each 0-frame codon (X-coordinate) within the nascent polypeptide chain on the mKate: eGFP intensity ratio as determined by DMS. Substitutions are arranged with respect to the relative abundance of tRNA that can decode each codon in HEK293 cells (see Materials and Methods). Ratios were normalized relative to that of WT, and black squares indicate mutations that lack sufficient coverage. The positions of TM3 (blue), the preceeding 6K residues (orange), and the slip-site (yellow) are shown for reference. The residues that are likely to reside within the exit tunnel or outside of the ribosome during ribosomal frameshifting are indicated with dashed lines.

post-translational modification sites) as well as the features in the transcript that stimulate −1PRF (slip-site and stem–loop). Retaining these features allowed us to identify and map effectors in a manner that cannot be achieved using a conventional dual-luciferase reporter system bearing an N-terminal fusion domain (Supplementary Figure S9). We should note that this reporter contains a proteolytic cleavage site that releases the E2 protein, which contains the first 55 mutagenized positions (residues 696–751), from the TF protein that is genetically fused to the mKate frameshift reporter. Thus, the observed trends within TMs 1 & 2 are unlikely to be sensitive to any changes in E2 stability. However, mutations downstream of this cleavage site could potentially alter the stability of the reporter protein in a way that affects the observed −1PRF signal. Nevertheless, we suspect such instances must be rare given that the observed trends within the slip-site and stem–loop region are highly consistent with previous findings (Figure 2A). Mutagenic effects within the heptanucleotide slip-site scale with the number of modified bases (Figure 2B), and the most sensitive bases downstream of the slip-site correspond to those that stabilize a previously characterized stem–loop (Figure 2C) (25,26). Beyond the validation of previous investigations, these measurements also show that −1PRF can be enhanced by mutations that strengthen or extend this stem–loop (Figure 2A). The fact that the stability of this stem–loop has not evolved to maximize −1PRF provides a clear

example of how SINV has evolved to maintain a specific frameshifting efficiency.

We recently showed that ribosomal frameshifting in the SINV structural polyprotein is sensitive to mutations that perturb the translocon-mediated membrane integration of TM2 (22). Considering the marginal hydrophobicity of this segment, we initially expected mutagenic trends within this region to reveal a simple relationship between the topological energetics of TM2 and the efficiency of −1PRF. However, the observed mutagenic trends cannot be explained by hydrophobicity alone (Supplementary Figure S5). Instead, our DMS and molecular dynamics data suggest that stimulatory pulling forces are specifically generated by the transfer of the C-terminal portion of TM2 from the translocon to membrane. At this final stage of translocation, these residues interact with the lumenal edge of the lateral gate, which pulls the loop between TMs 2 and 3 into the protein conducting channel of the translocon (Figure 5). This loop is wedged between TM2 and the translocon at the point of frameshifting, and the observed mutagenic patterns within this region do not simply track with hydrophobicity or codon usage (Figures 3 and 6). This observation suggests this segment may form specific interactions that influence membrane integration. Nevertheless, it is unclear whether mutations within this loop alter −1PRF by reducing the force on the nascent chain or by delaying the transmission of the force to the peptidyl transferase cen-
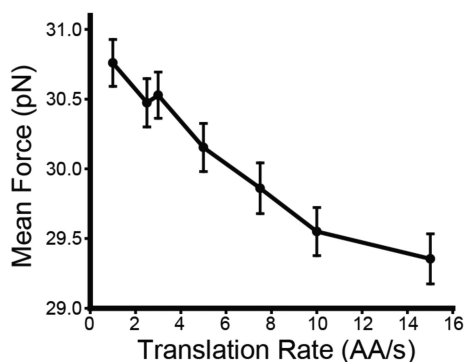
Figure 7. Impact of translation kinetics on nascent chain pulling forces. Coarse-grained molecular dynamics simulations of the translocon-mediated cotranslational folding of the SINV structural polyprotein were carried out varying the translation rate. Pulling forces on the nascent chain were recorded over 0.8 s while the ribosome occupies the slip-site. The mean force on the nascent chain across 100 independent trajectories is plotted against the corresponding translation rate. Error bars reflect the standard error of the mean.

ter until after the ribosome passes the slip-site. It is also unclear how these forces impact the decoding of the slip-site and the efficiency of ribosomal frameshifting. These forces are potentially propagated back to the catalytic core in a manner that influences translation kinetics (50–52), which is critical for the efficiency of −1PRF (15–19). Additional investigations are needed to explore how this network of interactions between the nascent chain, translocon, and lipid bilayer ultimately impact the efficiency of translational recoding.

Nascent polypeptides begin to fold during the early stages of translation (53,54). However, both cotranslational interactions between the nascent chain and ribosome and the steric restrictions imposed by the ribosomal exit tunnel tend to destabilize native protein structures (55,56). We were therefore surprised to find that −1PRF is highly sensitive to mutations within the region encoding the portion of the nascent chain that occupies the exit tunnel during frameshifting (Figure 3). Nevertheless, a recent investigation of the sequence constraints in other PRF motifs found that frameshifting is often enhanced by rare codons upstream of the slip-site (24). Our analysis of the effects of single-codon mutations suggests changes in −1PRF may be related to the effects of mutations on the relative abundance of decoding tRNAs (Figure 6) and CGMD simulations suggest stimulatory forces are sensitive to the rate of translation (Figure 7). These observations suggest the coupling between folding and −1PRF is likely kinetically controlled. Recent simulations suggest the forces generated by cotranslational folding are tied to the relative rates of translation and cotranslational folding (49). Translation rates could therefore represent a key determinant for the mechanical coupling between frameshifting and cotranslational folding. However, it remains unclear how much the single-codon mutations reported herein actually impact the physiological rate of translation and the estimated impact of translation kinetics on pulling forces appears to be relatively modest in relation to the magnitude of their effects on −1PRF. Given

these caveats, it may be that these mutations impact −1PRF in more than one way. For instance, it is possible that the slow translation within this region is critical to provide the time needed for the stem–loop to refold before the next ribosome reaches the slip-site. The native codon bias in this region may also control the frequency of ribosomal collisions, which were recently found to stimulate −1PRF (57). Modifications to the structure of the nascent chain could also potentially be relevant to −1PRF. This region encodes TM3, and it was recently found that hydrophobic helices exhibit a tendency to form helical structure within the exit tunnel (58). Furthermore, native −1PRF levels are only maintained when mutations introduce other hydrophobic side chains within this segment (Figure 3). Thus, the formation of helical structure within the exit tunnel could potentially also contribute to this mechanochemical coupling. Taken together, it seems the mechanistic basis of the mutagenic effects within this region are undoubtedly complex, and additional investigations are needed to tease apart each of these variables.

Together, our results reveal that the net efficiency of −1PRF in the alphavirus structural polyprotein arises from the coupling between multiple structural effectors. While this motif features the canonical slip-site and stem loop architecture found in most −1PRF sites, its efficiency is rendered tunable by mutations that alter mechanochemical transitions within the nascent chain and/ or the translation kinetics upstream of the slip-site. Our DMS data demonstrate that this degeneracy effectively expands the pool of genetic variations that are capable of adjusting the efficiency of −1PRF; an important consideration for viral evolution. Though this investigation focuses on a mechanochemical transition within the ribosome-translocon complex, it seems likely that many other classes of conformational transitions within the nascent chain may also be capable of stimulating −1PRF. Further investigations are needed to evaluate whether similar mechanisms are operative within other transcripts.

DATA AVAILABILITY

Flow cytometry data are available on FlowRepository (https://flowrepository.org/id/FR-FCM-Z3MG). Sequencing data are available in an NCBI BioProject (https://www.ncbi.nlm.nih.gov/bioproject/720960). Computational scripts for DMS are available on GitHub (https://github.com/schebachlab/DMS-analysis-Carmody). Computational scripts for coarse-grained molecular dynamics simulations of translocon-mediated cotranslational folding are available on GitHub (https://github.com/mhzimmer1/SecCGMD). The structural model of the translocation intermediate is available on ModelArchive (https://www.modelarchive.org/doi/10.5452/ma-n2gxk, Access Code: ohZTM7KNWb). Any other data and/ or analyses will be made freely available by the corresponding authors upon request.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

REFERENCES

1. Atkins,J.F., Loughran,G., Bhatt,P.R., Firth,A.E. and Baranov,P.V. (2016) Ribosomal frameshifting and transcriptional slippage: From genetic steganography and cryptography to adventitious use. *Nucleic Acids Res.*, **44**, 7007–7078.
2. Penn,W.D., Harrington,H.R., Schlebach,J.P. and Mukhopadhyay,S. (2020) Regulators of viral frameshifting: more than RNA influences translation events. *Annu. Rev. Virol.*, **7**, 219–238.
3. Hallengard,D., Kakoulidou,M., Lulla,A., Kummerer,B.M., Johansson,D.X., Mutso,M., Lulla,V., Fazakerley,J.K., Roques,P., Le Grand,R. *et al.* (2014) Novel attenuated Chikungunya vaccine candidates elicit protective immunity in C57BL/6 mice. *J. Virol.*, **88**, 2858–2866.
4. Rogers,K.J., Jones-Burrage,S., Maury,W. and Mukhopadhyay,S. (2020) TF protein of Sindbis virus antagonizes host type I interferon responses in a palmitoylation-dependent manner. *Virology*, **542**, 63–70.
5. Snyder,J.E., Kulcsar,K.A., Schultz,K.L., Riley,C.P., Neary,J.T., Marr,S., Jose,J., Griffin,D.E. and Kuhn,R.J. (2013) Functional characterization of the alphavirus TF protein. *J. Virol.*, **87**, 8511–8523.
6. Taylor,A., Melton,J.V., Herrero,L.J., Thaa,B., Karo-Astover,L., Gage,P.W., Nelson,M.A., Sheng,K.C., Lidbury,B.A., Ewart,G.D. *et al.* (2016) Effects of an in-frame deletion of the 6k gene locus from the genome of Ross river virus. *J. Virol.*, **90**, 4150–4159.
7. Napthine,S., Ling,R., Finch,L.K., Jones,J.D., Bell,S., Brierley,I. and Firth,A.E. (2017) Protein-directed ribosomal frameshifting temporally regulates gene expression. *Nat. Commun.*, **8**, 15582.
8. Wang,X., Xuan,Y., Han,Y., Ding,X., Ye,K., Yang,F., Gao,P., Goff,S.P. and Gao,G. (2019) Regulation of HIV-1 Gag-Pol expression by shiftless, an inhibitor of programmed -1 ribosomal frameshifting. *Cell*, **176**, 625–635.
9. Jacks,T., Madhani,H.D., Masiarz,F.R. and Varmus,H.E. (1988) Signals for ribosomal frameshifting in the Rous sarcoma virus gag-pol region. *Cell*, **55**, 447–458.
10. Bock,L.V., Caliskan,N., Korniy,N., Peske,F., Rodnina,M.V. and Grubmuller,H. (2019) Thermodynamic control of -1 programmed ribosomal frameshifting. *Nat. Commun.*, **10**, 4598.
11. Reil,H., Kollmus,H., Weidle,U.H. and Hauser,H. (1993) A heptanucleotide sequence mediates ribosomal frameshifting in mammalian cells. *J. Virol.*, **67**, 5579–5584.
12. Stahl,G., McCarty,G.P. and Farabaugh,P.J. (2002) Ribosome structure: revisiting the connection between translational accuracy and unconventional decoding. *Trends Biochem. Sci.*, **27**, 178–183.
13. Giedroc,D.P. and Cornish,P.V. (2009) Frameshifting RNA pseudoknots: structure and mechanism. *Virus Res.*, **139**, 193–208.
14. Mouzakis,K.D., Lang,A.L., Vander Meulen,K.A., Easterday,P.D. and Butcher,S.E. (2013) HIV-1 frameshift efficiency is primarily determined by the stability of base pairs positioned at the mRNA entrance channel of the ribosome. *Nucleic Acids Res.*, **41**, 1901–1913.
15. Caliskan,N., Katunin,V.I., Belardinelli,R., Peske,F. and Rodnina,M.V. (2014) Programmed -1 frameshifting by kinetic partitioning during impeded translocation. *Cell*, **157**, 1619–1631.
16. Chen,J., Petrov,A., Johansson,M., Tsai,A., O'Leary,S.E. and Puglisi,J.D. (2014) Dynamic pathways of -1 translational frameshifting. *Nature*, **512**, 328–332.
17. Choi,J., O'Loughlin,S., Atkins,J.F. and Puglisi,J.D. (2020) The energy landscape of -1 ribosomal frameshifting. *Sci. Adv.*, **6**, eaax6969.
18. Kim,H.K., Liu,F., Fei,J., Bustamante,C., Gonzalez,R.L. Jr and Tinoco,I. Jr (2014) A frameshifting stimulatory stem loop destabilizes the hybrid state and impedes ribosomal translocation. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, 5538–5543.
19. Korniy,N., Samatova,E., Anokhina,M.M., Peske,F. and Rodnina,M.V. (2019) Mechanisms and biomedical implications of -1 programmed ribosome frameshifting on viral and bacterial mRNAs. *FEBS Lett.*, **593**, 1468–1482.
20. Meydan,S., Klepacki,D., Karthikeyan,S., Margus,T., Thomas,P., Jones,J.E., Khan,Y., Briggs,J., Dinman,J.D., Vázquez-Laslop,N. *et al.* (2017) Programmed ribosomal frameshifting generates a copper transporter and a copper chaperone from the same gene. *Mol. Cell*, **65**, 207–219.
21. Yordanova,M.M., Wu,C., Andreev,D.E., Sach,M.S. and Atkins,J.F. (2015) A nascent peptide signal responsive to endogenous levels of polyamines acts to stimulate regulatory frameshifting on antizyme mRNA. *J. Biol. Chem.*, **290**, 17863–17878.
22. Harrington,H.R., Zimmer,M.H., Chamness,L.M., Nash,V., Penn,W.D., Miller,T.F. 3rd, Mukhopadhyay,S. and Schlebach,J.P. (2020) Cotranslational folding stimulates programmed ribosomal frameshifting in the alphavirus structural polyprotein. *J. Biol. Chem.*, **295**, 6798–6808.
23. Bhatt,P.R., Scaiola,A., Loughran,G., Leibundgut,M., Kratzel,A., Meurs,R., Dreos,R., O'Connor,K. M., McMillan,A., Bode,J.W. *et al.* (2021) Structural basis of ribosomal frameshifting during translation of the SARS-CoV-2 RNA genome. *Science*, **372**, 1306–1313.
24. Mikl,M., Pilpel,Y. and Segal,E. (2020) High-throughput interrogation of programmed ribosomal frameshifting in human cells. *Nat. Commun.*, **11**, 3061.
25. Chung,B.Y., Firth,A.E. and Atkins,J.F. (2010) Frameshifting in alphaviruses: a diversity of 3' stimulatory structures. *J. Mol. Biol.*, **397**, 448–456.
26. Kutchko,K.M., Madden,E.A., Morrison,C., Plante,K.S., Sanders,W., Vincent,H.A., Cruz Cisneros,M.C., Long,K.M., Moorman,N.J., Heise,M.T. *et al.* (2018) Structural divergence creates new functional features in alphavirus genomes. *Nucleic Acids Res.*, **46**, 3657–3670.
27. Niesen,M.J.M., Zimmer,M.H. and Miller,T.F. 3rd (2020) Dynamics of co-translational membrane protein integration and translocation via the Sec translocon. *J. Am. Chem. Soc.*, **142**, 5449–5460.
28. Wrenbeck,E.E., Klesmith,J.R., Stapleton,J.A., Adeniran,A., Tyo,K.E. and Whitehead,T.A. (2016) Plasmid-based one-pot saturation mutagenesis. *Nat. Methods*, **13**, 928–930.
29. Kendra,J.A., Advani,V.M., Chen,B., Briggs,J.W., Zhu,J., Bress,H.J., Pathy,S.M. and Dinman,J.D. (2018) Functional and structural characterization of the chikungunya virus translational recoding signals. *J. Biol. Chem.*, **293**, 17536–17545.
30. Khan,Y.A., Loughran,G. and Atkins,J. F. (2019) Contesting the evidence for −1 frameshifting in immune-functioning C-C chemokine receptor 5 (CCR5) – the HIV-1 co-receptor. bioRxiv doi: https://doi.org/10.1101/513333, 08 January 2019, preprint: not peer reviewed.
31. Matreyek,K.A., Stephany,J.J. and Fowler,D.M. (2017) A platform for functional assessment of large variant libraries in mammalian cells. *Nucleic Acids Res.*, **45**, e102.
32. Penn,W.D., McKee,A.G., Kuntz,C.P., Woods,H., Nash,V., Gruenhagen,T.C., Roushar,F.J., Chandak,M., Hemmerich,C., Rusch,D.B. *et al.* (2020) Probing biophysical sequence constraints within the transmembrane domains of rhodopsin by deep mutational scanning. *Sci. Adv.*, **6**, eaay7505.
33. Kirchner,S., Cai,Z., Rauscher,R., Kastelic,N., Anding,M., Czech,A., Kleizen,B., Ostedgaard,L.S., Braakman,I., Sheppard,D.N. *et al.* (2017) Alteration of protein function by a silent polymorphism linked to tRNA abundance. *PLoS Biol.*, **15**, e2000779.

34. Jacobs,J.L. and Dinman,J.D. (2004) Systematic analysis of bicistronic reporter assay data. *Nucleic Acids Res.*, **32**, e160.

35. Hessa,T., Meindl-Beinker,N.M., Bernsel,A., Kim,H., Sato,Y., Lerch-Bader,M., Nilsson,I., White,S.H. and Heijne,G. (2007) Molecular code for transmembrane-helix recognition by the Sec61 translocon. *Nature*, **450**, 1026–1030.

36. Niesen,M.J., Wang,C.Y., Van Lehn,R.C. and Miller,T.F. 3rd (2017) Structurally detailed coarse-grained model for Sec-facilitated co-translational protein translocation and membrane integration. *PLoS Comput. Biol.*, **13**, e1005427.

37. Niesen,M.J.M., Muller-Lucks,A., Hedman,R., von Heijne,G. and Miller,T.F. 3rd (2018) Forces on nascent polypeptides during membrane insertion and translocation via the Sec translocon. *Biophys. J.*, **115**, 1885–1894.

38. Voorhees,R.M., Fernandez,I.S., Scheres,S.H. and Hegde,R.S. (2014) Structure of the mammalian ribosome-Sec61 complex to 3.4 A resolution. *Cell*, **157**, 1632–1643.

39. Ma,C., Wu,X., Sun,D., Park,E., Catipovic,M.A., Rapoport,T.A., Gao,N. and Li,L. (2019) Structure of the substrate-engaged SecA-SecY protein translocation machine. *Nat. Commun.*, **10**, 2872.

40. Webb,B. and Sali,A. (2016) Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinformatics*, **54**, 5.6.1–5.6.37.

41. Bowers,K., Chow,E., Xu,H., Dror,R.O. and Eastwood,M.P. (2006) Scalable algorithms for molecular dynamics simulations on commodity clusters. *ACM/IEEE Supercomp. Conf.*, **6**, 7695–7700.

42. Chiasson,M.A., Rollins,N.J., Stephany,J.J., Sitko,K.A., Matreyek,K.A., Verby,M., Sun,S., Roth,F.P., DeSloover,D., Marks,D.S. *et al.* (2020) Multiplexed measurement of variant abundance and activity reveals VKOR topology, active site and human variant impact. *Elife*, **9**, e58026.

43. Brierley,I., Jenner,A.J. and Inglis,S.C. (1992) Mutational analysis of the "slippery-sequence" component of a coronavirus ribosomal frameshifting signal. *J. Mol. Biol.*, **227**, 463–479.

44. Yan,S., Wen,J.D., Bustamante,C. and Tinoco,I. Jr (2015) Ribosome excursions during mRNA translocation mediate broad branching of frameshift pathways. *Cell*, **160**, 870–881.

45. Korniy,N., Goyal,A., Hoffmann,M., Samatova,E., Peske,F., Pohlmann,S. and Rodnina,M.V. (2019) Modulation of HIV-1 Gag/Gag-Pol frameshifting by tRNA abundance. *Nucleic Acids Res.*, **47**, 5210–5222.

46. Halma,M.T.J., Ritchie,D.B., Cappellano,T.R., Neupane,K. and Woodside,M.T. (2019) Complex dynamics under tension in a high-efficiency frameshift stimulatory structure. *Proc. Natl. Acad. Sci. U.S.A.*, **116**, 19500–19505.

47. Ritchie,D.B., Foster,D.A. and Woodside,M.T. (2012) 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.

48. Moon,C.P. and Fleming,K.G. (2011) Side-chain hydrophobicity scale derived from transmembrane protein folding into lipid bilayers. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 10174–10177.

49. Leininger,S.E., Trovato,F., Nissley,D.A. and O'Brien,E.P. (2019) Domain topology, stability, and translation speed determine mechanical force generation on the ribosome. *Proc. Natl. Acad. Sci. U.S.A.*, **116**, 5523–5532.

50. Fritch,B., Kosolapov,A., Hudson,P., Nissley,D.A., Woodcock,H.L., Deutsch,C. and O'Brien,E.P. (2018) Origins of the mechanochemical coupling of peptide bond formation to protein synthesis. *J. Am. Chem. Soc.*, **140**, 5077–5087.

51. Jiang,Y. and O'Brien,E.P. (2021) Mechanical forces have a range of effects on the rate of ribosome catalyzed peptidyl transfer depending on direction. *J. Phys. Chem. B*, **125**, 7128–7136.

52. Leininger,S.E., Narayan,K., Deutsch,C. and O'Brien,E.P. (2019) Mechanochemistry in translation. *Biochemistry*, **58**, 4657–4666.

53. Kosolapov,A. and Deutsch,C. (2009) Tertiary interactions within the ribosomal exit tunnel. *Nat. Struct. Mol. Biol.*, **16**, 405–411.

54. O'Brien,E.P., Hsu,S.D., Christodoulou,J., Vendruscolo,M. and Dobson,C.M. (2010) Transient tertiary structure formation within the ribosome exit port. *J. Am. Chem. Soc.*, **132**, 16928–16937.

55. Kudva,R., Tian,P., Pardo-Avila,F., Carroni,M., Best,R.B., Bernstein,H.D. and von Heijne,G. (2018) The shape of the bacterial ribosome exit tunnel affects cotranslational protein folding. *Elife*, **7**, e36326.

56. Samelson,A.J., Jensen,M.K., Soto,R.A., Cate,J.H. and Marqusee,S. (2016) Quantitative determination of ribosome nascent chain stability. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, 13402–13407.

57. Smith,A.M., Costello,M.S., Kettring,A.H., Wingo,R.J. and Moore,S.D. (2019) Ribosome collisions alter frameshifting at translational reprogramming motifs in bacterial mRNAs. *Proc. Natl. Acad. Sci. U.S.A.*, **116**, 21769–21779.

58. Bano-Polo,M., Baeza-Delgado,C., Tamborero,S., Hazel,A., Grau,B., Nilsson,I., Whitley,P., Gumbart,J.C., von Heijne,G. and Mingarro,I. (2018) Transmembrane but not soluble helices fold inside the ribosome tunnel. *Nat. Commun.*, **9**, 5246.