

---

# Conformational dynamics of the frameshift stimulatory structure in HIV-1

---

DUSTIN B. RITCHIE,<sup>1</sup> TONIA R. CAPPELLANO,<sup>1</sup> COLLIN TITTLE,<sup>1</sup> NEGAR REZAJOOEI,<sup>1</sup> LOGAN ROULEAU,<sup>1</sup> WILLIAM K.A. SIKKEMA,<sup>1</sup> and MICHAEL T. WOODSIDE<sup>1,2</sup>

<sup>1</sup>Department of Physics, University of Alberta, Edmonton AB T6G 2E1, Canada

<sup>2</sup>National Institute for Nanotechnology, National Research Council, Edmonton AB T6G 2M9, Canada

## ABSTRACT

Programmed ribosomal frameshifting (PRF) in HIV-1 is thought to be stimulated by a hairpin in the mRNA, although a pseudoknot-like triplex has also been proposed. Because the conformational dynamics of the stimulatory structure under tension applied by the ribosomal helicase during translation may play an important role in PRF, we used optical tweezers to apply tension to the HIV stimulatory structure and monitor its unfolding and refolding dynamics. The folding and unfolding kinetics and energy landscape of the hairpin were measured by ramping the force on the hairpin up and down, providing a detailed biophysical characterization. Unexpectedly, whereas unfolding reflected the simple two-state behavior typical of many hairpins, refolding was more complex, displaying significant heterogeneity. Evidence was found for multiple refolding pathways as well as previously unsuspected, partially folded intermediates. Measuring a variant mRNA containing only the sequence required to form the proposed triplex, it behaved largely in the same way. Nonetheless, very rarely, high-force unfolding events characteristic of pseudoknot-like structures were observed. The rare occurrence of the triplex suggests that the hairpin is the functional stimulatory structure. The unusual heterogeneity of the hairpin dynamics under tension suggests a possible functional role in PRF similar to the dynamics of other stimulatory structures.

**Keywords:** programmed ribosomal frameshifting; RNA folding; force spectroscopy; optical tweezers

## INTRODUCTION

In programmed ribosomal frameshifting (PRF), the ribosome is induced to shift into a different reading frame by specific signals in the messenger RNA (mRNA), typically bypassing the original stop codon and continuing protein synthesis in the new reading frame (Giedroc and Cornish 2009; Brierley et al. 2010). PRF into the  $-1$  reading frame is used by many viruses, including HIV, West Nile Virus, and SARS coronavirus as well as plant viruses and certain bacteriophages, to produce multiple protein products from a single mRNA. Despite decades of research, PRF mechanisms are still incompletely understood (Giedroc and Cornish 2009; Brierley et al. 2010; Dinman 2012; Caliskan et al. 2015). PRF is stimulated by a stable structure formed by the viral mRNA downstream from a “slippery sequence” where the frameshift actually occurs. Together, these elements of the frameshift signal achieve a stringent regulation of  $-1$  PRF levels, which is essential for virus propagation (Dinman and Wickner 1992; Dulude et al. 2006; Plant et al. 2010). However, the details of how PRF is induced

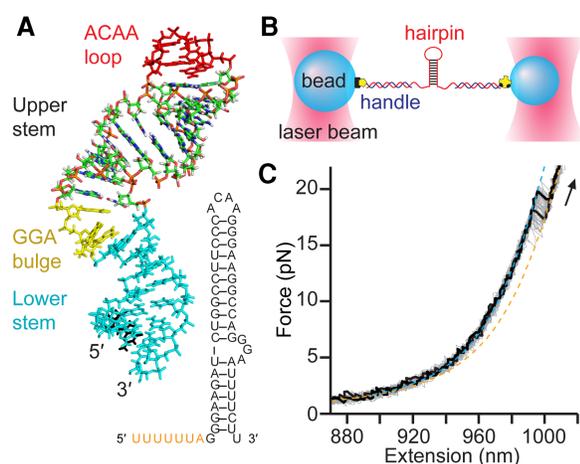
and its level regulated by frameshift signals remain controversial, and many different models have been proposed (Brierley et al. 2010). A feature common to several of these models is that the mRNA is under tension while the ribosome attempts to resolve the stimulatory structure during the critical step for frameshifting, suggesting that the properties of the stimulatory structure under mechanical tension may play an important role.

This view has motivated the study of the mechanical properties of stimulatory structures, using force probes like optical tweezers to perform single-molecule force spectroscopy (SMFS). SMFS, wherein force is applied to a molecule and the molecular extension measured as the structure changes in response, is a powerful tool for probing the conformational dynamics of single molecules (Ritchie and Woodside 2015), exploring questions like the role of intermediate states and competing pathways (Onoa et al. 2003; Chen et al. 2007; Li et al. 2007; Greenleaf et al. 2008; Stigler et al. 2011; Frieda and Block 2012; Gao et al. 2012; Ritchie et al. 2012; Yu et al.

© 2017 Ritchie et al. This article is distributed exclusively by the RNA Society for the first 12 months after the full-issue publication date (see <http://rnajournal.cshlp.org/site/misc/terms.xhtml>). After 12 months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

---

Corresponding author: [michael.woodside@ualberta.ca](mailto:michael.woodside@ualberta.ca)  
Article is online at <http://www.rnajournal.org/cgi/doi/10.1261/rna.061655.117>.



**FIGURE 1.** Structure and force spectroscopy of the full-length HIV-1 frameshift-stimulating hairpin. (A) The three-dimensional NMR structure of the full-length HIV-1 frameshift-stimulating hairpin (Staple and Butcher 2005) is shown with key structural features indicated. The secondary structure is shown with the slippery sequence (orange). (B) Experimental setup consisting of a single RNA hairpin molecule tethered between two beads held in optical traps. (C) Individual unfolding FECs (black) are plotted above the aggregated data from 50 FECs (gray). The FECs show a monotonic rise of force with extension until the RNA contour length changes abruptly upon unfolding, causing a “rip” in the FEC. Worm-like chain (WLC) fits to the folded (cyan) and unfolded (orange) branches of the FECs yield the contour length change upon unfolding. Arrow denotes direction of force ramp.

2012, 2015) and measuring the energy landscape underlying the behavior (Woodside and Block 2014). Pseudoknots, formed when nucleotides within a hairpin loop base pair with complementary nucleotides outside of that loop (Dam et al. 1992), are the most common PRF stimulatory structure. Initial suggestions that  $-1$  PRF levels are determined by the resistance of the pseudoknot to mechanical unfolding (Hansen et al. 2007; Chen et al. 2009) were found not to hold true when surveying a wider range of pseudoknots (Ritchie et al. 2012). Instead, the conformational plasticity of the pseudoknot under tension has been suggested to play an important role (Ritchie et al. 2012, 2014; de Messieres et al. 2014). Such studies have not yet been extended to other types of stimulatory structure, however, to test whether conformational plasticity plays a similar role.

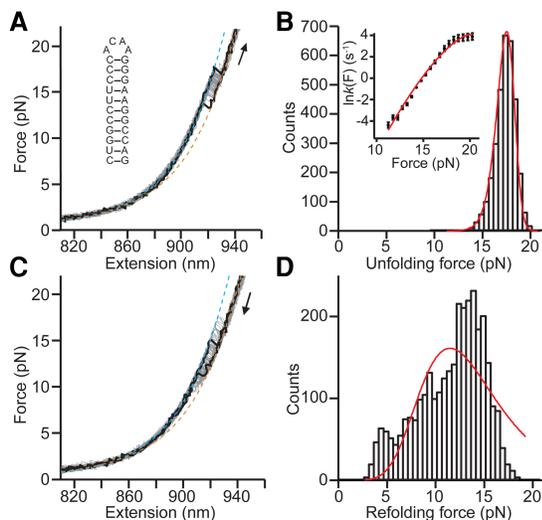
HIV-1 is unusual among eukaryotic viruses in that the stimulatory structure is thought to be a simple hairpin, as opposed to a pseudoknot. A structure for this hairpin has been solved by NMR (Fig. 1A), featuring a two-stem helix containing a 3-purine bulge separating an 11-bp stable upper stem and a weak, U-rich lower stem (Dulude et al. 2002; Gaudin et al. 2005; Staple and Butcher 2005). However, the lower stem of the hairpin must be unwound and its 5' end sequestered inside the ribosome during frameshifting in order to position the slippery site properly. The notion that the upper stem alone is the stimulatory structure (Jacks et al. 1988; Wilson et al. 1988; Parkin et al. 1992; Kollmus et al. 1996; Bidou et al. 1997) is supported by the observation that it

can indeed induce significant PRF levels (Wilson et al. 1988; Parkin et al. 1992; Bidou et al. 1997). An alternative model has been proposed, however, whereby a pseudoknot-like triplex composed of interactions between the upper stem and the 3' arm of the lower stem forms upon lower-stem unwinding by the ribosome (Dinman et al. 2002; Hong et al. 2010; Huang et al. 2014).

Previous work on the mechanical unfolding of the HIV-1 stimulatory hairpin using optical tweezers found that the upper stem was mechanically stable but not the lower stem (Mazaauric et al. 2009). Neither the heterogeneity of the dynamics nor the energy landscape was investigated, however. We revisited the HIV-1 frameshift signal to explore its properties under tension in more detail, finding that the upper stem hairpin displayed heterogeneous refolding dynamics unlike those typically observed for simple stem-loop structures, suggesting possible parallels with the conformational plasticity seen in stimulatory pseudoknots. We also tested whether the proposed pseudoknot-like triplex occurs by using a truncated version of the full-length hairpin sequence, finding that although the triplex can form, it does so sufficiently rarely that it is unlikely to be the functional stimulatory structure.

## RESULTS

We first investigated the mechanical unfolding of the full-length HIV-1 frameshifting hairpin, containing both the stable upper stem and U-rich lower stem (Fig. 1A). Constructs containing the mRNA hairpin flanked on each side by kb-long DNA-RNA duplex handles were attached to beads held in optical traps (Fig. 1B). The traps were held close together for 3 sec to permit zero-force folding of the RNA, then separated at constant velocity to ramp up the force applied to the RNA while the molecular extension was measured, generating force-extension curves (FECs). Single molecules were unfolded and refolded repeatedly by increasing and decreasing the force, resulting in a distribution of FECs displaying the range of possible behaviors. FECs for unfolding the full-length hairpin (Fig. 1C) showed a characteristic nonlinear rise of force with extension as the handles were stretched, interrupted by an abrupt extension increase and concomitant force decrease as the hairpin unfolded cooperatively (Liphardt et al. 2001; Woodside et al. 2008). A single discrete unfolding transition was typically observed in the range 14–21 pN (122 FECs), although sometimes multiple transitions were seen. Fitting the FECs to extensible worm-like chain (WLC) models (Fig. 1C, dotted lines), the contour length change upon unfolding,  $\Delta L_c$ , was found to be consistent with unfolding only the upper stem:  $\Delta L_c = 13 \pm 1$  nm (all errors represent S.E.M.), in agreement with previous work (Mazaauric et al. 2009) and the expectation from the NMR structure ( $\Delta L_c = 13.5$  nm). The lower stem unfolding occurred at very low force, as found previously (Mazaauric et al. 2009), making it difficult to observe directly.



**FIGURE 2.** Force spectroscopy of the HIV-1 upper stem alone. (A) Individual unfolding FECs (black) are plotted *above* the aggregated data from 100 FECs (gray). WLC fits are shown for the folded (cyan) and unfolded (orange) branches of the FECs. (*Inset*) Secondary structure of the HIV-1 upper stem alone. (B) The distribution of unfolding forces for HIV-1 upper stem alone FECs (black) is fit well by Equation 2 (red), yielding parameters describing the mechanical resistance to unfolding. (*Inset*) Force-dependent unfolding rate (black) is fit well by Equation 3 (red). (C) Individual refolding FECs (black) are plotted along with the aggregated data from 100 FECs (gray). WLC fits are shown for the unfolded (orange) and folded (cyan) branches of the FECs. (D) The distribution of refolding forces for HIV-1 upper stem alone FECs (black) is poorly fit by Equation 2 (red), a consequence of heterogeneous refolding that is more complicated than the single-barrier two-state behavior observed for unfolding.

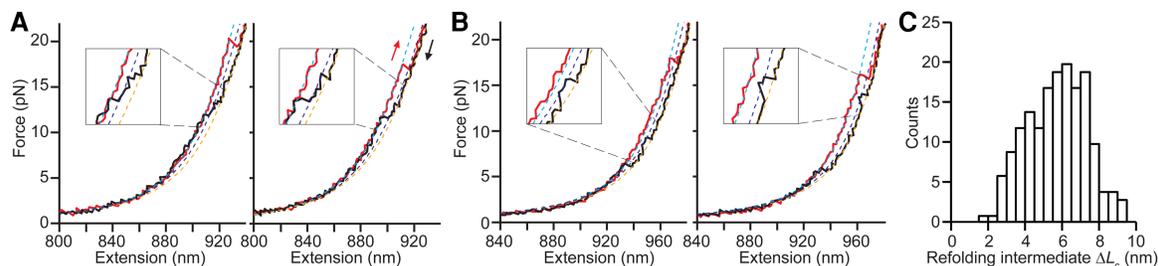
Because the lower stem is almost certainly unfolded during PRF in order to satisfy spacer-length requirements (Léger et al. 2004; Staple and Butcher 2005), we focused primarily on the behavior of the upper stem alone. Unfolding FEC measurements were repeated for a construct consisting of the upper stem alone (Fig. 2A, inset). Discrete unfolding events in the FECs (Fig. 2A) were observed in the same force range as for the full construct, 14–20 pN, and with the same  $\Delta L_c$  from WLC fits (Fig. 2A, dashed lines),  $12.5 \pm 0.5$  nm, confirming that the discrete unfolding transition seen for the full-length hairpin did indeed involve only the upper stem. Such two-state behavior is very typical for simple stem-loop structures under tension (Liphardt et al. 2001; Manosas et al. 2006; Woodside et al. 2006a,b). Investigating the unfolding of the upper stem in more detail, we examined the distribution of unfolding forces in the FECs (Fig. 2B, black),  $p(F)$ , to quantify the resistance of the hairpin to mechanical unfolding. The average unfolding force from 3280 FECs, here  $17 \pm 1$  pN, provided the simplest measure of the mechanical stability. Other parameters, describing the unfolding rate and the shape of the underlying energy landscape, were obtained by fitting  $p(F)$  to Equation 2 (Fig. 2B, red). As seen,  $p(F)$  was exactly the shape expected for a two-state system with a single barrier (Dudko et al. 2006)

and was well-fit by Equation 2. A complementary analysis of the kinetics based on the cumulative probability of unfolding (Dudko et al. 2008) was also used to obtain the unfolding rate as a function of force,  $k(F)$  (Fig. 2B, inset), which was again well fit by the same type of landscape model using Equation 3. These fits yielded  $\log k_{\text{off}} [\text{s}^{-1}] = -9.9 \pm 0.7$ ,  $\Delta x^\ddagger = 8.2 \pm 0.8$  nm, and  $\Delta G^\ddagger = 38 \pm 4 k_B T$ .

Turning to the refolding FECs (Fig. 2C), discrete refolding events were observed that looked qualitatively similar to the unfolding events but occurred over a lower force range (3–18 pN). WLC fitting revealed that  $\Delta L_c$  during refolding,  $13.0 \pm 0.7$  nm, was the same (within error) as for unfolding, indicating that the hairpin was switching between the same two states. However, the refolding force distribution (Fig. 2D) was qualitatively different: Where one would normally expect to see a smooth rise to the most probable unfolding force followed by a swift fall to zero as the force was reduced (the mirror image of the distribution for the unfolding forces), there was instead a fat tail in the distribution at low forces, and possibly several discrete peaks. Proceeding nevertheless to fit the refolding  $p(F)$  (Fig. 2D, red) from 3277 refolding FECs to the analog of Equation 2 for refolding force distributions (Pierse and Dudko 2013), we found  $\log k_{\text{on}} [\text{s}^{-1}] = 3.3 \pm 0.5$ ,  $\Delta x^\ddagger = 1.3 \pm 0.3$  nm,  $\Delta G^\ddagger = 5 \pm 3 k_B T$ . However, the fit was not particularly good and thus these results are likely not reliable—indeed the shape of  $p(F)$  for refolding suggests that refolding is more complex than a single-barrier, two-state system. Fat tails or multiple peaks in the force distribution have been shown to be a consequence of heterogeneity in the folding behavior (Hyeon et al. 2014; Pierse and Dudko 2017).

Examining the refolding curves in more detail provided additional support for the view that refolding was more complex than unfolding. Roughly 5% of curves showed evidence of partially folded intermediate states. In some cases, the intermediate state was on-pathway between the unfolded and folded states, as seen by the progression from unfolded to intermediate to folded (Fig. 3A). Notably, such intermediates were never seen in the unfolding curves, indicating that new pathways between the unfolded and folded states were being observed during refolding that were not observed during unfolding. In other cases, the intermediates appeared as transient excursions out of the unfolded state that returned to the unfolded state before the folded structure was attained (Fig. 3B), suggesting that they may be off-pathway states that represent yet another behavior not seen during unfolding. The distribution of contour length changes for the intermediates, measured relative to the unfolded state via WLC fits (Fig. 3A,B insets, dotted lines), was quite broad (Fig. 3C), indicating the presence of several different intermediate structures.

In order to explore the possible formation of the proposed pseudoknot-like triplex, we next investigated a construct, denoted  $\Delta 5'FL$ , in which the 5' arm of the lower stem in the full-length hairpin was deleted to mimic the mRNA sequence that



**FIGURE 3.** Refolding intermediates in FECs for the upper stem alone. (A) Representative FECs with on-pathway refolding intermediates showing individual unfolding (red) and refolding (black) FECs with WLC fits to the folded (cyan) and unfolded (orange) branches and the intermediate state (blue). The folding progresses from unfolded to folded through the intermediate, indicating the latter is on-pathway. (B) Representative FECs with off-pathway refolding intermediates showing individual unfolding (red) and refolding (black) FECs with WLC fits to the folded (cyan) and unfolded (orange) branches and the intermediate state (blue). The intermediate state appears as a transient excursion out of the unfolded state that returns to the unfolded state before the folded structure is attained, indicating that it is likely off-pathway. (C) Distribution of contour length changes from the intermediate states to the unfolded state. The broad range indicates the presence of several different intermediate structures.

would be free to fold after unwinding of the lower stem by the ribosome (Fig. 4A, inset). Based on 2714 FECs, the unfolding of  $\Delta 5'$ FL was almost always identical to that of the upper stem alone (Fig. 4A, black). WLC fits of the unfolding FECs found  $\Delta L_c = 13 \pm 1$  nm, and the average unfolding force was  $18 \pm 1$  pN, the same in each case (within error) as for upper-stem unfolding. The distribution of unfolding forces (Fig. 4B) was well fit by Equation 2, and the parameters found from fitting  $p(F)$  to Equation 2 (Fig. 4B, red) and  $k(F)$  to Equation 3 (Fig. 4B, inset) were in excellent agreement with those obtained for the upper stem alone:  $\log k_{\text{off}} [\text{s}^{-1}] = -9.7 \pm 0.9$ ,  $\Delta x^\ddagger = 7.4 \pm 0.7$  nm, and  $\Delta G^\ddagger = 39 \pm 6 k_B T$ .

Turning to the refolding of  $\Delta 5'$ FL, we found that refolding FECs, too, were very similar to those for upper stem only (Fig. 4C). The same  $\Delta L_c$  (within error) was found as for the upper stem,  $13 \pm 1$  nm from 2713 FECs. Furthermore, a similarly broad distribution of refolding forces was observed (Fig. 4D), and intermediates were again detected in a small subset of refolding FECs (Fig. 4E), with a range of contour lengths similar to that for the upper stem, again reflective of extensive heterogeneity in the refolding dynamics. The qualitative and quantitative similarities between the unfolding and refolding of the upper stem only and the  $\Delta 5'$ FL construct lead us to conclude that the  $\Delta 5'$ FL sequence must form the upper-stem hairpin structure almost all the time.

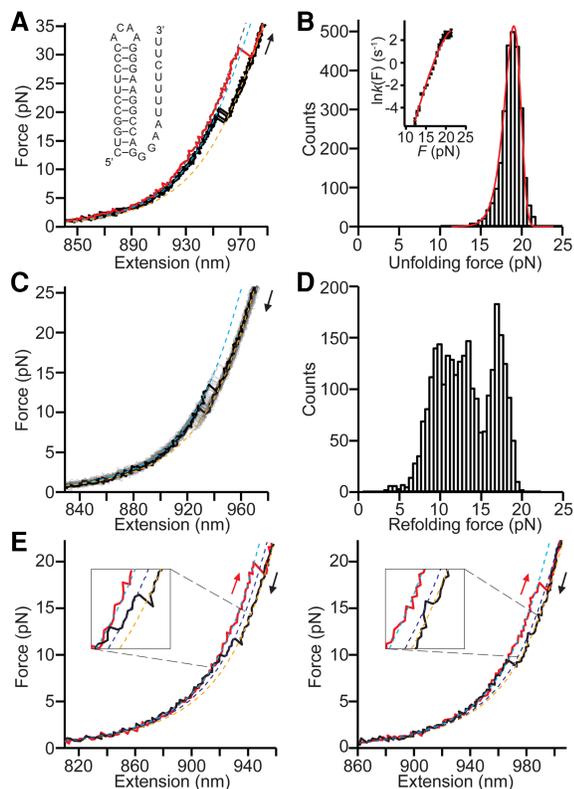
In a handful of cases ( $\sim 0.1\%$  of the pulls), however, we saw unfolding FECs that did not fit this pattern (Fig. 4A, red). These curves displayed a much higher unfolding force, roughly 30 pN as opposed to 18 pN, in conjunction with noticeably larger contour length change,  $16 \pm 1$  nm versus  $13 \pm 1$  nm. These differences are both significant and revealing. A larger  $\Delta L_c$  than for the upper stem is only possible if nucleotides that are not part of the upper-stem hairpin (i.e., those in the 3' arm of the lower stem) are participating in this new structure. Based on the proposed triplex structure (Fig. 4A; Dinman et al. 2002), we would expect  $\Delta L_c \sim 17 \pm 1$  nm, in good agreement with the observed value. The high unfolding force can only be explained by the presence of tertiary struc-

ture: It is much higher than the unfolding forces observed previously for nucleic acid secondary structure (Liphardt et al. 2001; Li et al. 2006; Manosas et al. 2006; Woodside et al. 2006b; Chen et al. 2007; Green et al. 2008; Greenleaf et al. 2008; Mazauric et al. 2009; Neupane et al. 2011; Bercy and Bockelmann 2015). Instead, it is similar to the typical unfolding force of about 30–40 pN observed for most pseudoknots (Chen et al. 2007; Green et al. 2008; Ritchie et al. 2012, 2014; de Messieres et al. 2014; Zhong et al. 2016), suggesting that the  $\Delta 5'$ FL mRNA can indeed form a pseudoknot-like structure as proposed, although it does so only very rarely.

## DISCUSSION

These results provide insight into how the frameshift signal may stimulate  $-1$  PRF in HIV-1. Addressing first the question of whether frameshifting may be stimulated by a putative triplex structure, we found that a structure consistent with this pseudoknot-like triplex can indeed form when the 5' strand of the lower stem is removed, mimicking the situation when this strand is sequestered under the ribosome footprint just before the frameshift occurs. However, the notion that the frameshift signal partitions into a triplex structure that actively triggers  $-1$  PRF and a hairpin that is inactive in PRF is contradicted by the observation that the triplex structure formed roughly two orders of magnitude less frequently than would be needed to account for the typical 5% PRF efficiency in HIV-1 (at least under our measurement conditions). Instead, the similarity between the behavior of the  $5'\Delta$ FL and upper-stem constructs across all the characteristics examined (length changes and forces for unfolding and refolding, intermediates and heterogeneity during refolding) supports the view that the upper-stem hairpin alone is the stimulatory structure.

Given that most stimulatory structures are pseudoknots, not hairpins (Giedroc and Cornish 2009), a key question is what property of this hairpin accounts for its unusual ability to trigger  $-1$  PRF efficiently, in contrast to other hairpins that



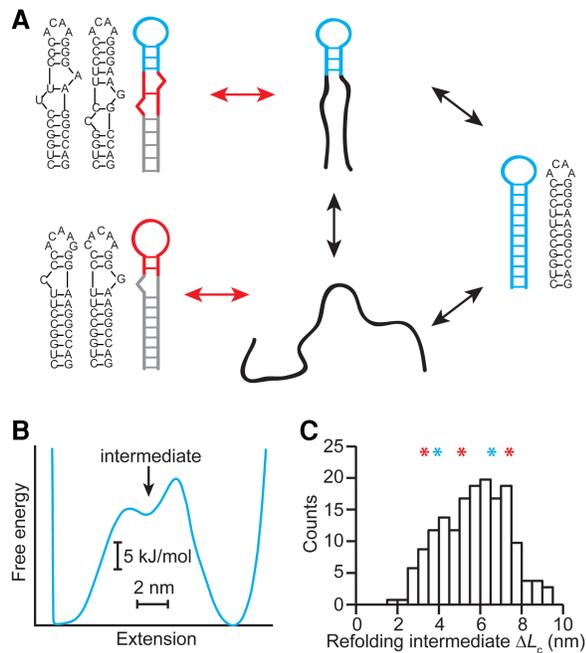
**FIGURE 4.** Force spectroscopy of the  $\Delta 5'$ FL construct. (A) Most FECs show length changes and unfolding forces corresponding to those expected for unfolding the upper stem alone (black). A few FECs show a longer length and a higher unfolding force, indicating the presence of a larger structure containing tertiary contacts (red). (Dotted lines) WLC fits to the unfolded state (orange), folded upper stem (cyan), and folded triplex state (blue). (*Inset*) Secondary structure of the  $\Delta 5'$ FL construct. (B) The unfolding force distribution for the events with shorter  $\Delta L_c$  (black) is well fit by Equation 2 (red), yielding the same fitting parameters as for the upper stem alone. (*Inset*) Force-dependent unfolding rate (black) fit by Equation 3 (red). (C) Refolding FECs (black, gray) all have  $\Delta L_c$  consistent with the upper stem only. (Dotted lines) WLC fits to the unfolded state (orange) and folded upper stem (cyan). (D) The refolding force distribution is again broad, a consequence of heterogeneous refolding, similar to that for the upper stem alone. (E) Some refolding FECs (black) contain intermediate states not observed in unfolding FECs (red). A broad range of intermediate-state  $\Delta L_c$  values from WLC fits (orange: unfolded, blue: intermediate, cyan: folded) indicates the presence of multiple intermediate states, as for the upper stem only.

are poor PRF stimulators. The unfolding force, barrier height and compliance, and unfolding rate are all different from pseudoknots, but these properties have previously been shown to be unrelated to PRF stimulation efficiency (Ritchie et al. 2012). More intriguing as an explanation is the heterogeneity in the refolding: The HIV-1 stimulatory hairpin has folding dynamics that are unusually complex for a simple duplex stem-loop, which typically unfold and refold as two-state systems (Liphardt et al. 2001; Woodside et al. 2006a,b) along a single pathway (Neupane et al. 2015, 2016) in a way that can be predicted from the sequence (Manosas et al. 2006; Woodside et al. 2006b; Neupane

et al. 2015). The HIV-1 hairpin, in contrast, shows evidence of multiple refolding pathways likely involving both on- and off-pathway intermediates, behavior typically associated with more complex structures including bulges or helix junctions as in molecules like the HIV-1 transactivation response (TAR) hairpin (Li et al. 2006) or purine riboswitch aptamers (Neupane et al. 2011). Such heterogeneity is not unheard-of in simple stem-loops, as both temperature-jump measurements (Ma et al. 2006) and computational simulations (Bowman et al. 2008) suggested the presence of multiple folding pathways in the folding of hairpins with highly structured tetraloops similar to the loop in the HIV-1 hairpin, but it has not previously been suggested to play a functional role.

We note that the broad refolding force distributions are consistent with the presence of both static and dynamic heterogeneity, arising from conformational fluctuations that occur on timescales that are, respectively, much slower than or similar to the timescale for refolding (Hyeon et al. 2014). Static heterogeneity would result in a distribution containing multiple peaks (Pierse and Dudko 2017), each the characteristic shape for a single refolding transition (Pierse and Dudko 2013), whereas dynamic heterogeneity produces a distinct tail at low forces (Hyeon et al. 2014). The observed distributions provide evidence for both types of heterogeneity: Multiple peaks may be discerned, but there is also a low-force tail. Because it is difficult to separate the static and dynamic components reliably, however, we are unable to quantify the heterogeneity by fitting the distributions. Nevertheless, we can test the picture we propose to account for our observations, involving refolding via native or nonnative loops and on- or off-pathway intermediates (Fig. 5A), by modeling the intermediates. The existence of an on-pathway intermediate is supported by modeling of the energy landscape expected from the sequence (Woodside et al. 2006a), which shows a shallow well halfway between the folded and unfolded states (Fig. 5B). Structure-prediction tools like *mfold* (Zuker 2003) also reveal several possible off-pathway intermediates having native or nonnative loops (Fig. 5, insets). The range of length changes associated with these intermediates,  $\Delta L_c \sim 3\text{--}7.5$  nm, is fully consistent with the values observed in the refolding FECs (Fig. 5C), supporting the picture in Figure 5A.

The heterogeneity observed in the refolding of the HIV stimulatory hairpin evokes a key feature of pseudoknot folding that has been linked to efficient PRF—the fact that PRF efficiency correlates well with the conformational plasticity of the stimulatory structure (Ritchie et al. 2012, 2014). This connection may help to explain how this hairpin can act as a stimulatory structure in  $-1$  PRF in lieu of a more typical pseudoknot: Similar to the way that frameshift-inducing pseudoknots differ from other pseudoknots in having more heterogeneous dynamics (reflected in a greater tendency to form alternate structures when refolding), so too does the HIV frameshifting hairpin display more heterogeneous dynamics than expected for a simple hairpin structure. We



**FIGURE 5.** Model of heterogeneous folding for HIV-1 upper stem alone. (A) The folding is proposed to include two-state folding between the unfolded (black) and native (cyan) states, as well as on-pathway intermediates containing a native loop (cyan) but an incomplete stem, and off-pathway intermediates having nonnative stems or loops (red) that must be resolved before the native structure can form (all suboptimal structure predictions from *mfold*). (B) A model of the energy landscape for native folding under tension (Woodside et al. 2006a) shows a shallow well that could support an intermediate state. (C) The  $\Delta L_c$  values expected from the proposed intermediate states (asterisks: red—off-pathway, cyan—on-pathway) match the range of values in the measured histogram (black).

speculated previously that such heterogeneity might play a role via fluctuations in the tension in the mRNA generated by structural fluctuations in the stimulatory structure induced by the mechanical tension applied by the ribosome, fluctuations that are then communicated to the tRNA binding sites to induce  $-1$  PRF (Ritchie et al. 2012). Our results for the HIV stimulatory hairpin are consistent with this picture, suggesting that it extends beyond pseudoknots and may be a general feature of structures that are efficient at stimulating PRF. A role for structural fluctuations in triggering  $-1$  PRF has also been suggested for a hairpin-based stimulatory structure in an alphavirus (Kendra et al. 2017).

It is important to note that hairpin-stimulated frameshifting has not been studied as extensively as PRF stimulated by pseudoknots, and diverse stimulatory structures may not necessarily induce frameshifting via identical mechanisms. Some studies of hairpin-induced frameshifting have led to the proposal that hairpins stimulate PRF by acting as a simple roadblock, but robust quantitative correlations remain elusive (Bidou et al. 1997; Yu et al. 2011; Mouzakis et al. 2013) and such a picture is inconsistent with the properties of pseudoknots (Ritchie et al. 2012, 2014). It remains unclear

precisely how dynamics in the stimulatory structure (not to mention other elements like interactions between the stimulatory structure and the ribosome) lead to  $-1$  PRF, and much work remains to be done before a complete mechanistic model of  $-1$  PRF can be built. Nevertheless, identifying features like conformational heterogeneity that are shared between diverse frameshift signals provide important clues that may help clarify this long-standing problem.

## MATERIALS AND METHODS

### Sample preparation and measurement

Sequences encoding the full-length hairpin, the upper stem only, or the full-length hairpin with 5' strand of the lower stem deleted were inserted into the pLuc-1 plasmid between the SpeI and BamHI restriction sites. The resulting transcription template, containing a hairpin flanked by linker regions on either side, was amplified by PCR and transcribed *in vitro* using T7 RNA polymerase. Two ssDNA handles (one complementary to 840 nt on the 3' end of the transcript and labeled with biotin, the other complementary to 2280 nt on the 5' end of the transcript and labeled with digoxigenin) were produced by asymmetric PCR from double-stranded DNA PCR products corresponding to the flanking handle sequences (Neupane et al. 2011). The handles were annealed with the RNA transcript, then incubated with 600 and 820 nm diameter polystyrene beads labeled with avidin DN (Vector Labs) and anti-digoxigenin (Roche), respectively, to create dumbbells. Dumbbells were placed in measuring buffer (50 mM MOPS, pH 7.0, 130 mM KCl, 4 mM MgCl<sub>2</sub>, 50 U/mL Superase•In RNase inhibitor [Ambion] and oxygen scavenging system: 40 U/mL glucose oxidase, 185 U/mL catalase, and 8.3 mg/mL glucose) and inserted into a sample chamber on a clean microscope slide in the optical trap.

Force-extension curves (FECs) were measured with a custom-built, dual-beam optical trap similar to one described previously (Neupane et al. 2011). Briefly, two orthogonally polarized beams from a single 1064-nm laser were steered independently with acousto-optic deflectors to create two traps. The motion of beads held in the traps was detected by collecting the light from two orthogonally polarized 830-nm laser beams aligned on the traps that was scattered by the beads onto position-sensitive diodes. Trap stiffnesses were 0.58 and 0.37 pN/nm, calibrated as described previously (Svoboda and Block 1994); each RNA construct was measured using a different stiffness calibration, possibly giving rise to systematic differences of up to 5%–10%. The traps were moved apart at a constant speed of 110–270 nm/sec, sampling data at 20 kHz and filtering online at 10 kHz with an 8-pole Bessel filter.

### FEC analysis

To determine the contour length change during unfolding and refolding transitions,  $\Delta L_c$ , FECs were aligned to correct for small amounts of drift in the measurement and fit to extensible worm-like chains (WLCs) (Wang et al. 1997):

$$F(x) = \frac{k_B T}{L_p} \left[ \frac{1}{4} \left( 1 - \frac{x}{L_c} + \frac{F}{K} \right)^{-2} - \frac{1}{4} + \frac{x}{L_c} - \frac{F}{K} \right], \quad (1)$$

where  $L_c$  is the contour length of the polymer,  $L_p$  the persistence length,  $K$  the enthalpic elasticity, and  $k_B T$  the thermal energy. Two WLCs in series were used, one for the duplex handle ( $L_c \sim 1100$  nm,  $L_p \sim 35$  nm,  $K \sim 1200$  pN) and the other for the unfolded mRNA ( $L_c = 0.59$  nm/nt,  $L_p = 1$  nm,  $K = 2000$  pN). The change in RNA contour length was found by fitting the two branches of the FECs on either side of a transition. The contour length change expected from the NMR structure of the HIV-1 frameshift hairpin (Staple and Butcher 2003) was calculated using  $\Delta L_c = n_{nt} \cdot L_c^{nt} - d_T$ , where  $d_T$  is the distance between the termini of the folded hairpin as measured from the NMR structure (1.8 nm),  $n_{nt}$  is the number of nucleotides in the hairpin, and  $L_c^{nt} = 0.59$  nm/nt is the contour length per nucleotide (Saenger 1984). For the proposed HIV triplex structure,  $d_T$  was estimated from pseudoknots of similar size as 5 nm, yielding an expected  $\Delta L_c$  of  $17 \pm 1$  nm.

The distribution of unfolding forces was analyzed using the kinetic theory of Dudko et al. (2006) to determine the parameters characterizing the energy landscape underlying the unfolding behavior:

$$p(F) \propto \frac{k(F)}{r} \exp \left\{ \frac{k_{\text{off}}}{\beta \Delta x^\ddagger r} - \frac{k(F)}{\beta \Delta x^\ddagger r} \left( 1 - \frac{\Delta x^\ddagger F}{\Delta G^\ddagger} \nu \right)^{1-1/\nu} \right\}, \quad (2)$$

where

$$k(F) = k_{\text{off}} \left( 1 - \frac{\Delta x^\ddagger F}{\Delta G^\ddagger} \nu \right)^{1/\nu-1} \exp \left\{ \frac{\Delta G^\ddagger}{k_B T} \left[ 1 - \left( 1 - \frac{\Delta x^\ddagger F}{\Delta G^\ddagger} \nu \right)^{1/\nu} \right] \right\}, \quad (3)$$

$k_{\text{off}}$  is the unfolding rate at zero force,  $\Delta x^\ddagger$  is the distance to the transition state from the folded state,  $\Delta G^\ddagger$  is the height of the energy barrier,  $r$  is the loading rate,  $\beta$  is the inverse thermal energy, and  $\nu$  is a parameter characterizing the shape of the energy barrier ( $\nu = 1/2$  for a sharp cusp-like barrier,  $2/3$  for a softer linear-cubic potential). We averaged the results obtained under these two limiting cases for the shape of the energy barrier, since this shape was unknown.

The free energy change between the folded and unfolded states of the HIV upper stem was determined from the FECs using the Crooks fluctuation theorem (Crooks 1999; Collin et al. 2005). The work distributions for unfolding and refolding, respectively  $P_U(W)$  and  $P_F(W)$ , were calculated from integrals of the pulling curves, and the free-energy change was found from

$$\frac{P_U(W)}{P_F(-W)} = \exp(\beta[W - \Delta G]), \quad (4)$$

using the Bennett acceptance ratio to determine the intersection of  $P_U(W)$  and  $P_F(W)$  (Collin et al. 2005). The result,  $\Delta G = 23 \pm 1 k_B T$ , was compared with previous ensemble measurements of the upper stem using UV absorbance, which found  $\Delta G = 25 \pm 1 k_B T$  (Mouzakis et al. 2013), in reasonable agreement given the systematic uncertainty in the force calibration.

No significant changes in any of the properties of the RNA constructs were observed over time during repeated unfolding–refolding cycles. Molecule-to-molecule variations were small, with the average unfolding force varying by  $\sim 10\%$ , similar to the amount expected from calibration errors. Different numbers of pulling cycles were measured on different molecules, but this did not affect the qualitative analysis of the refolding heterogeneity: Resampling the refolding force distributions to ensure equal weighting from

each molecule produced histograms in which the range of forces and overall shape were not materially changed.

## ACKNOWLEDGMENTS

This work was supported by the Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council Canada, Alberta Innovates Technology Futures, and National Research Council Canada.

Received April 12, 2017; accepted May 12, 2017.

## REFERENCES

- Bercy M, Bockelmann U. 2015. Hairpins under tension: RNA versus DNA. *Nucleic Acids Res* **43**: 9928–9936.
- Bidou L, Stahl G, Grima B, Liu H, Cassan M, Rousset JP. 1997. In vivo HIV-1 frameshifting efficiency is directly related to the stability of the stem-loop stimulatory signal. *RNA* **3**: 1153–1158.
- Bowman GR, Huang X, Yao Y, Sun J, Carlsson G, Guibas LJ, Pande VS. 2008. Structural insight into RNA hairpin folding intermediates. *J Am Chem Soc* **130**: 9676–9678.
- Brierley I, Gilbert RJC, Pennell S. 2010. Pseudoknot-dependent programmed  $-1$  ribosomal frameshifting: structures, mechanisms and models. In *Recoding: expansion of decoding rules enriches gene expression* (ed. Atkins JF, Gesteland RF), *Nucleic Acids and Molecular Biology*, Vol. 24, pp. 149–174. Springer, New York.
- Caliskan N, Peske F, Rodnina MV. 2015. Changed in translation: mRNA recoding by  $-1$  programmed ribosomal frameshifting. *Trends Biochem Sci* **40**: 265–274.
- Chen G, Wen J-D, Tinoco I. 2007. Single-molecule mechanical unfolding and folding of a pseudoknot in human telomerase RNA. *RNA* **13**: 2175–2188.
- Chen G, Chang K-Y, Chou M-Y, Bustamante C, Tinoco I. 2009. Triplex structures in an RNA pseudoknot enhance mechanical stability and increase efficiency of  $-1$  ribosomal frameshifting. *Proc Natl Acad Sci* **106**: 12706–12711.
- Collin D, Ritort F, Jarzynski C, Smith SB, Tinoco I, Bustamante C. 2005. Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies. *Nature* **437**: 231–234.
- Crooks GE. 1999. Entropy production fluctuation theorem and the non-equilibrium work relation for free energy differences. *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics* **60**: 2721–2726.
- Dam E, Pleij K, Draper D. 1992. Structural and functional aspects of RNA pseudoknots. *Biochemistry* **31**: 11665–11676.
- de Messieres M, Chang J-C, Belew AT, Meskauskas A, Dinman JD, La Porta A. 2014. Single-molecule measurements of the CCR5 mRNA unfolding pathways. *Biophys J* **106**: 244–252.
- Dinman JD. 2012. Mechanisms and implications of programmed translational frameshifting. *Wiley Interdiscip Rev RNA* **3**: 661–673.
- Dinman JD, Wickner RB. 1992. Ribosomal frameshifting efficiency and gag/gag-pol ratio are critical for yeast M1 double-stranded RNA virus propagation. *J Virol* **66**: 3669–3676.
- Dinman JD, Richter S, Plant EP, Taylor RC, Hammell AB, Rana TM. 2002. The frameshift signal of HIV-1 involves a potential intramolecular triplex RNA structure. *Proc Natl Acad Sci* **99**: 5331–5336.
- Dudko OK, Hummer G, Szabo A. 2006. Intrinsic rates and activation free energies from single-molecule pulling experiments. *Phys Rev Lett* **96**: 108101.
- Dudko OK, Hummer G, Szabo A. 2008. Theory, analysis, and interpretation of single-molecule force spectroscopy experiments. *Proc Natl Acad Sci* **105**: 15755–15760.
- Dulude D, Baril M, Brakier-Gingras L. 2002. Characterization of the frameshift stimulatory signal controlling a programmed  $-1$  ribosomal frameshift in the human immunodeficiency virus type 1. *Nucleic Acids Res* **30**: 5094–5102.

- Dulude D, Berchiche YA, Gendron K, Brakier-Gingras L, Heveker N. 2006. Decreasing the frameshift efficiency translates into an equivalent reduction of the replication of the human immunodeficiency virus type 1. *Virology* **345**: 127–136.
- Frieda KL, Block SM. 2012. Direct observation of cotranscriptional folding in an adenine riboswitch. *Science* **338**: 397–400.
- Gao Y, Zorman S, Gundersen G, Xi Z, Ma L, Sirinakis G, Rothman JE, Zhang Y. 2012. Single reconstituted neuronal SNARE complexes zipper in three distinct stages. *Science* **337**: 1340–1343.
- Gaudin C, Mazauric M-H, Traïkia M, Guittet E, Yoshizawa S, Fourmy D. 2005. Structure of the RNA signal essential for translational frameshifting in HIV-1. *J Mol Biol* **349**: 1024–1035.
- Giedroc DP, Cornish PV. 2009. Frameshifting RNA pseudoknots: structure and mechanism. *Virus Res* **139**: 193–208.
- Green L, Kim C-H, Bustamante C, Tinoco I Jr. 2008. Characterization of the mechanical unfolding of RNA pseudoknots. *J Mol Biol* **375**: 511–528.
- Greenleaf WJ, Frieda KL, Foster DAN, Woodside MT, Block SM. 2008. Direct observation of hierarchical folding in single riboswitch aptamers. *Science* **319**: 630–633.
- Hansen TM, Reihani SNS, Oddershede LB, Sørensen MA. 2007. Correlation between mechanical strength of messenger RNA pseudoknots and ribosomal frameshifting. *Proc Natl Acad Sci* **104**: 5830–5835.
- Hong H, Kim Y-G, Hohng S. 2010. Single-molecule FRET studies on frameshifting RNA structures of human immunodeficiency virus. *Bull Korean Chem Soc* **31**: 1021–1024.
- Huang X, Yang Y, Wang G, Cheng Q, Du Z. 2014. Highly conserved RNA pseudoknots at the Gag-Pol junction of HIV-1 suggest a novel mechanism of -1 ribosomal frameshifting. *RNA* **20**: 587–593.
- Hyeon C, Hinczewski M, Thirumalai D. 2014. Evidence of disorder in biological molecules from single molecule pulling experiments. *Phys Rev Lett* **112**: 138101.
- Jacks T, Power MD, Masiarz FR, Luciw PA, Barr PJ, Varmus HE. 1988. Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. *Nature* **331**: 280–283.
- Kendra JA, de la Fuente C, Brahm A, Woodson C, Bell TM, Chen B, Khan YA, Jacobs JL, Kehn-Hall K, Dinman JD. 2017. Ablation of programmed -1 ribosomal frameshifting in Venezuelan equine encephalitis virus results in attenuated neuropathogenicity. *J Virol* **91**: e01766-16.
- Kollmus H, Hentze MW, Hauser H. 1996. Regulated ribosomal frameshifting by an RNA-protein interaction. *RNA* **2**: 316–323.
- Léger M, Sidani S, Brakier-Gingras L. 2004. A reassessment of the response of the bacterial ribosome to the frameshift stimulatory signal of the human immunodeficiency virus type 1. *RNA* **10**: 1225–1235.
- Li PTX, Collin D, Smith SB, Bustamante C, Tinoco I Jr. 2006. Probing the mechanical folding kinetics of TAR RNA by hopping, force-jump, and force-ramp methods. *Biophys J* **90**: 250–260.
- Li PTX, Bustamante C, Tinoco I Jr. 2007. Real-time control of the energy landscape by force directs the folding of RNA molecules. *Proc Natl Acad Sci* **104**: 7039–7044.
- Liphardt J, Onoa B, Smith SB, Tinoco I Jr, Bustamante C. 2001. Reversible unfolding of single RNA molecules by mechanical force. *Science* **292**: 733–737.
- Ma H, Proctor DJ, Kierzek E, Kierzek R, Bevilacqua PC, Gruebele M. 2006. Exploring the energy landscape of a small RNA hairpin. *J Am Chem Soc* **128**: 1523–1530.
- Manosas M, Collin D, Ritort F. 2006. Force-dependent fragility in RNA hairpins. *Phys Rev Lett* **96**: 218301.
- Mazauric M-H, Seol Y, Yoshizawa S, Visscher K, Fourmy D. 2009. Interaction of the HIV-1 frameshift signal with the ribosome. *Nucleic Acids Res* **37**: 7654–7664.
- Mouzakis KD, Lang AL, Vander Meulen KA, Easterday PD, Butcher SE. 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.
- Neupane K, Yu H, Foster DAN, Wang F, Woodside MT. 2011. Single-molecule force spectroscopy of the add adenine riboswitch relates folding to regulatory mechanism. *Nucleic Acids Res* **39**: 7677–7687.
- Neupane K, Manuel AP, Lambert J, Woodside MT. 2015. Transition-path probability as a test of reaction-coordinate quality reveals DNA hairpin folding is a one-dimensional diffusive process. *J Phys Chem Lett* **6**: 1005–1010.
- Neupane K, Foster DAN, Dee DR, Yu H, Wang F, Woodside MT. 2016. Direct observation of transition paths during the folding of proteins and nucleic acids. *Science* **352**: 239–242.
- Onoa B, Dumont S, Liphardt J, Smith SB, Tinoco I Jr, Bustamante C. 2003. Identifying kinetic barriers to mechanical unfolding of the *T. thermophila* ribozyme. *Science* **299**: 1892–1895.
- Parkin NT, Chamorro M, Varmus HE. 1992. Human immunodeficiency virus type 1 gag-pol frameshifting is dependent on downstream mRNA secondary structure: demonstration by expression in vivo. *J Virol* **66**: 5147–5151.
- Pierce CA, Dudko OK. 2013. Kinetics and energetics of biomolecular folding and binding. *Biophys J* **105**: L19–L22.
- Pierce CA, Dudko OK. 2017. Distinguishing signatures of multipathway conformational transitions. *Phys Rev Lett* **118**: 088101.
- Plant EP, Rakauskaitė R, Taylor DR, Dinman JD. 2010. Achieving a golden mean: mechanisms by which coronaviruses ensure synthesis of the correct stoichiometric ratios of viral proteins. *J Virol* **84**: 4330–4340.
- Ritchie DB, Woodside MT. 2015. Probing the structural dynamics of proteins and nucleic acids with optical tweezers. *Curr Opin Struct Biol* **34**: 43–51.
- Ritchie DB, Foster DAN, Woodside MT. 2012. Programmed -1 frameshifting efficiency correlates with RNA pseudoknot conformational plasticity, not resistance to mechanical unfolding. *Proc Natl Acad Sci* **109**: 16167–16172.
- Ritchie DB, Soong J, Sikkema WKA, Woodside MT. 2014. Anti-frameshifting ligand reduces the conformational plasticity of the SARS virus pseudoknot. *J Am Chem Soc* **136**: 2196–2199.
- Saenger W. 1984. *Principles of nucleic acid structure*. Springer, New York.
- Staple DW, Butcher SE. 2003. Solution structure of the HIV-1 frameshift inducing stem-loop RNA. *Nucleic Acids Res* **31**: 4326–4331.
- Staple DW, Butcher SE. 2005. Solution structure and thermodynamic investigation of the HIV-1 frameshift inducing element. *J Mol Biol* **349**: 1011–1023.
- Stigler J, Ziegler F, Gieseke A, Gebhardt JCM, Rief M. 2011. The complex folding network of single calmodulin molecules. *Science* **334**: 512–516.
- Svoboda K, Block SM. 1994. Biological applications of optical forces. *Annu Rev Biophys Biomol Struct* **23**: 247–285.
- Wang MD, Yin H, Landick R, Gelles J, Block SM. 1997. Stretching DNA with optical tweezers. *Biophys J* **72**: 1335–1346.
- Wilson W, Braddock M, Adams SE, Rathjen PD, Kingsman SM, Kingsman AJ. 1988. HIV expression strategies: ribosomal frameshifting is directed by a short sequence in both mammalian and yeast systems. *Cell* **55**: 1159–1169.
- Woodside MT, Block SM. 2014. Reconstructing folding energy landscapes by single-molecule force spectroscopy. *Annu Rev Biophys* **43**: 19–39.
- Woodside MT, Anthony PC, Behnke-Parks WM, Larizadeh K, Herschlag D, Block SM. 2006a. Direct measurement of the full, sequence-dependent folding landscape of a nucleic acid. *Science* **314**: 1001–1004.
- Woodside MT, Behnke-Parks WM, Larizadeh K, Travers K, Herschlag D, Block SM. 2006b. Nanomechanical measurements of the sequence-dependent folding landscapes of single nucleic acid hairpins. *Proc Natl Acad Sci* **103**: 6190–6195.
- Woodside MT, García-García C, Block SM. 2008. Folding and unfolding single RNA molecules under tension. *Curr Opin Chem Biol* **12**: 640–646.

- Yu C-H, Noteborn MH, Pleij CWA, Olsthoorn RCL. 2011. Stem-loop structures can effectively substitute for an RNA pseudoknot in  $-1$  ribosomal frameshifting. *Nucleic Acids Res* **39**: 8952–8959.
- Yu H, Liu X, Neupane K, Gupta AN, Brigley AM, Solanki A, Sosova I, Woodside MT. 2012. Direct observation of multiple misfolding pathways in a single prion protein molecule. *Proc Natl Acad Sci* **109**: 5283–5288.
- Yu H, Dee DR, Liu X, Brigley AM, Sosova I, Woodside MT. 2015. Protein misfolding occurs by slow diffusion across multiple barriers in a rough energy landscape. *Proc Natl Acad Sci* **112**: 8308–8313.
- Zhong Z, Yang L, Zhang H, Shi J, Vandana JJ, Lam DTUH, Olsthoorn RCL, Lu L, Chen G. 2016. Mechanical unfolding kinetics of the SRV-1 gag-pro mRNA pseudoknot: possible implications for  $-1$  ribosomal frameshifting stimulation. *Sci Rep* **6**: 39549.
- Zuker M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* **31**: 3406–3415.