



Mechanical strength of RNA knot in Zika virus protects against cellular defenses

Meng Zhao and Michael T. Woodside

Unusual knot-like structures recently discovered in viral exoribonuclease-resistant RNAs (xrRNAs) prevent digestion by host RNases to create subgenomic RNAs enhancing infection and pathogenicity. xrRNAs are proposed to prevent digestion through mechanical resistance to unfolding. However, their unfolding force has not been measured, and the factors determining RNase resistance are unclear. Furthermore, how these knots fold remains unknown. Unfolding a Zika virus xrRNA with optical tweezers revealed that it was the most mechanically stable RNA yet observed. The knot formed by threading the 5' end into a three-helix junction before pseudoknot interactions closed a ring around it. The pseudoknot and tertiary contacts stabilizing the threaded 5' end were both required to generate extreme force resistance, whereas removing a 5'-end contact produced a low-force knot lacking RNase resistance. These results indicate mechanical resistance plays a central functional role, with the fraction of molecules forming extremely high-force knots determining the RNase resistance level.

Cells infected by viruses typically attempt to digest the viral RNA with RNase enzymes. By inserting structures that are resistant to RNase digestion into the viral genome, viruses can turn this defense back against the host cell, using host RNases to generate subgenomic RNAs in downstream protected regions^{1–4}. These exoribonuclease-resistant RNAs (xrRNAs) were first discovered in flaviviruses^{5,6}, the family of viruses including dengue, Zika, yellow fever and chikungunya that cause some of the most prevalent infections world-wide. However, they have since been identified in several other virus families, including plant viruses, suggesting that they are a general class of biologically active RNA important to viruses^{1,4,7–9}. They may therefore present a fruitful new target for antiviral therapeutics.

In flaviviruses such as Zika virus, xrRNAs occur in the upstream end of the 3' untranslated region and resist 5' → 3' degradation by Xrn1, a cytoplasmic host exoribonuclease that helps to ensure proper messenger RNA turnover and purging of aberrant RNA^{10,11}. In many cases, at least two xrRNAs are present in tandem, producing an array of noncoding subgenomic RNAs of different lengths from incomplete digestion of the viral RNA^{6,12}. These RNAs accumulate to high levels in infected cells and interact with various host proteins to affect processes including normal cellular RNA decay¹³, RNA interference¹⁴, interferon response¹⁵ and vector–host transmission^{16,17}, thereby increasing viral pathogenicity¹⁸. In some plant viruses, xrRNAs are found between open reading frame regions, suggesting xrRNAs may also regulate expression of viral proteins associated with propagation⁹.

Structural studies of xrRNAs^{1–4} have revealed a previously unknown fold architecture, featuring a single-stranded segment threaded through a ring consisting of a three-helix junction that is closed on one end by pseudoknot (PK) interactions to form a knot-like structure (Fig. 1a), which we term a ring knot. The ring-knot fold is proposed to be responsible for exoribonuclease resistance (XR) by blocking digestion mechanically, preventing the unfolding of xrRNAs by the helicase activity of RNases¹⁹. This hypothesis is supported by measurements of xrRNA translocation through nanopores, which found much higher resistance to translocation from the 5' end than the 3' end, mirroring the directionality of XR activity²⁰. However, the mechanical stability of xrRNAs has

not yet been quantified by measuring their unfolding force, and it is unclear what sets XR levels, including whether XR arises generically from the knot-like fold topology or if it relates to specific mechanical properties. Furthermore, although fluorescence measurements have probed the dynamics of ring closing¹, the full sequence of events leading to ring-knot folding—placing the formation of secondary-structure elements and tertiary contacts in their relevant contexts—has yet to be established. It thus remains unclear how the knot folds, which interactions are critical to the putative mechanical resistance and how these interactions relate to XR.

Here we address these questions through single-molecule force spectroscopy (SMFS) of an xrRNA from Zika virus. Using optical tweezers²¹ to apply force to the ends of single xrRNA molecules (Fig. 1b), we ramped the force up and down to induce unfolding and refolding, observing directly the complete sequence of steps followed. By measuring the extension of the xrRNA as its structure changed, we identified intermediates in the folding and deduced their structures. We found that the core of the native fold was very resistant to mechanical unfolding, but achieving this extreme resistance required a specific sequence of steps in the folding: formation first of a three-helix junction, then threading of the 5' end into a cleft in this junction and stabilization by tertiary contacts, before closure of the ring around the threaded 5' end by a PK. The formation of structures with lower mechanical resistance—whether knotted or unknotted—correlated strongly with sensitivity to exoribonuclease digestion for both wild-type and mutant xrRNAs, suggesting that extreme mechanical resistance is the critical feature conferring XR.

Results

Mechanical stability of wild-type Zika virus xrRNA. To study the mechanical properties of the xrRNA, RNA containing the xrRNA sequence (Fig. 1a and Supplementary Table 1) flanked on either side by kilobase-long handle sequences was annealed to complementary single-stranded DNA (handles attached to beads held in optical traps (Fig. 1b). The traps were then moved apart to ramp up the force on the RNA before being returned to ramp the force down, waiting 0.5–5.5 s at near-zero force for the xrRNA to refold before pulling again. This procedure generated force-extension

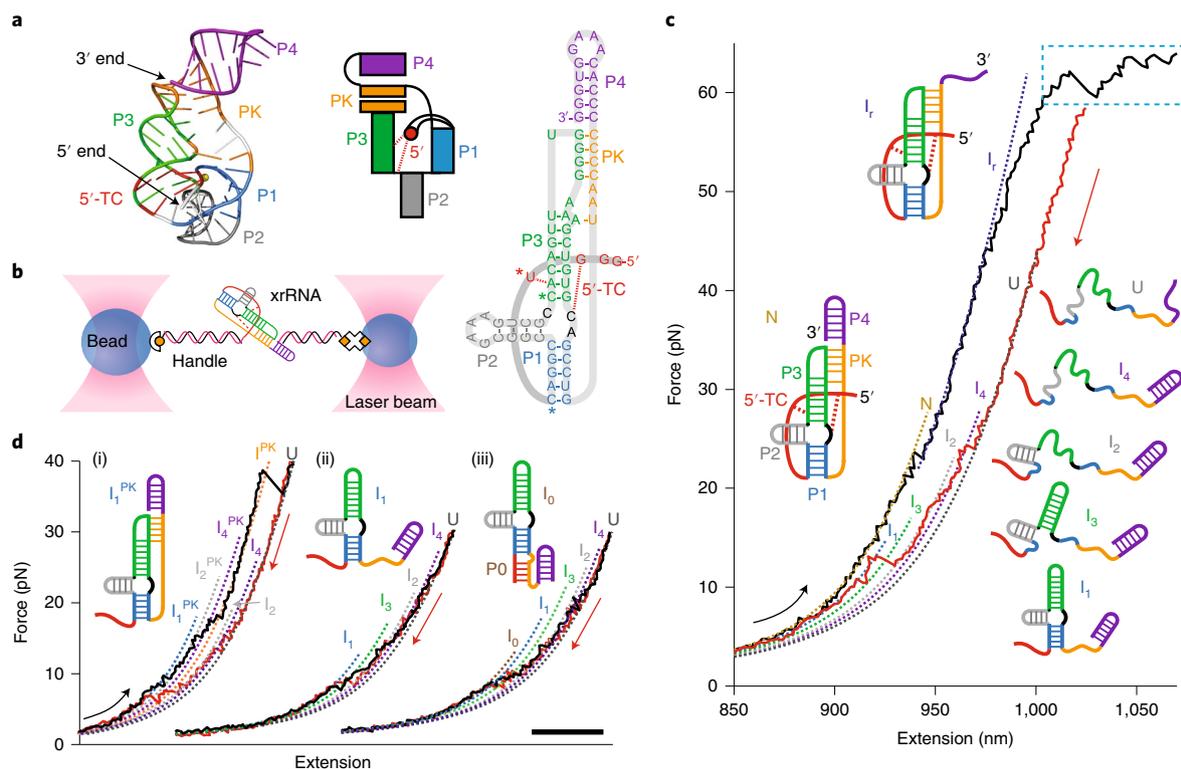


Fig. 1 | Single-molecule force spectroscopy of Zika virus xrRNA. **a**, Structure of the Zika virus xrRNA showing color-coded helices (P1–P4), PK and threading contacts (5′-TC). Left, crystal structure (PDB ID 5TPY) from ref. ³. Center, schematic of knot-like topology. Right, secondary structure (asterisks denote nucleotides coordinated with a Mg²⁺ ion). **b**, Measurement schematic showing a single xrRNA held under tension via DNA handles bound to beads held in optical traps. **c**, Most unfolding FECs (black) do not completely unfold until the overstretching transition (dashed box); refolding FECs preceding these unfolding FECs (red) show sequential formation of helices. Dashed lines show WLC fits of each state (native N, intermediates I_x , unfolded U). Insets show structural cartoons for each state, with color-coding as in **a**. **d**, The other types of FEC. (i) 4% of FECs show unfolding of an unthreaded PK structure (inset) at moderate force, roughly 30–50 pN. (ii) 15% of FECs involve secondary structure only (inset), with helices P1–P4 unfolding at forces of roughly 10–20 pN. (iii) 1% of FECs involve a secondary structure only but with a nonnative extension of helix P1 (P0, inset). Scale bar, 50 nm.

curves (FECs) reflecting the structural changes during unfolding and refolding: unfolding (refolding) of substructures within the RNA was detected as ‘rips’ in the curves (Fig. 1c,d) where the extension abruptly increased (decreased) concomitant with a decrease (increase) in the force.

Examining the unfolding of wild-type xrRNA in the presence of 4 mM Mg²⁺, most FECs (roughly 80%) showed a small rip at roughly 20 pN, but then typically remained featureless up to roughly 60 pN (Fig. 1c, black). Fitting the FECs before and after the low-force rip to extensible worm-like chain (WLC) models²² (Fig. 1c, dotted lines), we found a contour length change of $\Delta L_c = 6.1 \pm 0.4$ nm, only a small fraction of the 38.3 nm expected (Supplementary Table 2) for complete unfolding. The xrRNA thus remained in a mostly folded intermediate, denoted I_r , even at 60 pN—well above the unfolding forces reported for other compact RNA structures^{23–29} and in the same range as the duplex overstretching transition³⁰. Indeed, in most of these curves, unfolding the xrRNA completely required holding it at the overstretching force for at least 2 s; as a result, I_r unfolding occurred at the same time as handle overstretching and was usually difficult to distinguish (Fig. 1c, dashed box and Supplementary Fig. 1a). After waiting 2 s at roughly 65 pN, however, refolding curves displayed the full ΔL_c expected (Fig. 1c, red and Supplementary Fig. 1b), revealing multiple intermediates well fit by WLC models (Fig. 1c, dashed lines) that yielded the contour length changes for each state (Supplementary Table 2).

Turning to the unfolding FECs without extreme mechanical resistance (roughly 20% of all FECs), roughly 1/5 (4% of all FECs)

unfolded through three intermediate states at low forces (Fig. 1d(i) and Supplementary Fig. 1c), in the 10–20 pN range characteristic of secondary structures³¹, followed by a final transition at higher force, in the roughly 30–50 pN range characteristic of most tertiary structure unfolding^{23–27}. In these FECs, complete unfolding showed $\Delta L_c = 36 \pm 1$ nm, slightly shorter than for the native fold, indicating that the xrRNA started in a state close to, yet different from, the native fold. The remaining FECs without extreme mechanical resistance (16% of all FECs), in contrast, unfolded through low-force intermediates only, indicating no tertiary contacts. Most (15% of all FECs) had a total ΔL_c of 31.3 ± 0.8 nm (Fig. 1d(ii) and Supplementary Fig. 1d), shorter than expected and hence indicating the xrRNA was incompletely folded, whereas a very few (1% of all FECs) had an additional low-force transition yielding $\Delta L_c = 36.9 \pm 0.5$ nm (Fig. 1d(iii) and Supplementary Fig. 1e).

Multiple intermediate and nonnative states in xrRNA folding.

To interpret these results, we compared the ΔL_c values observed to the values expected for unfolding various substructures of the native fold (Supplementary Fig. 2 and Supplementary Table 2). For the FECs with extreme mechanical resistance (Fig. 1c), the low-force rip corresponded to unfolding of helix P4 (Fig. 1a, purple), leading to the structure for I_r illustrated in Fig. 1c. The refolding preceding such unfolding curves was consistent with sequential formation of first P4 (producing state I_4), then P2 (producing I_2), P3 (producing I_3), and finally P1 (producing I_1); the PK and 5′-end tertiary contacts (denoted 5′-TC) responsible for extreme unfolding forces

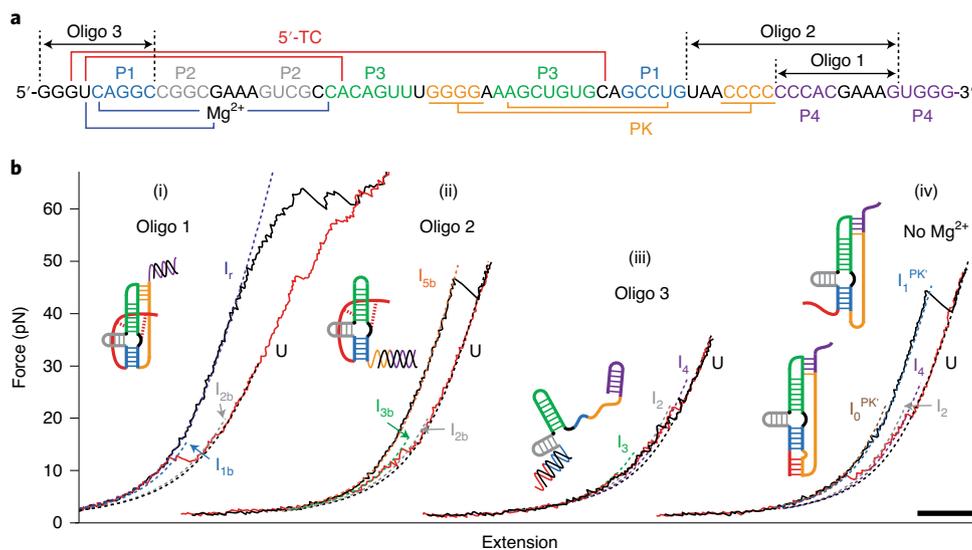


Fig. 2 | FECs in presence of antisense oligos. **a**, Locations where antisense oligos pair with xrRNA sequence: oligo 1 blocks formation of P4, oligo 2 blocks P4 and PK and oligo 3 blocks P1 and 5'-TC. **b**, Representative unfolding (black) and refolding (red) FECs in the presence of oligo 1 (i), oligo 2 (ii), oligo 3 (iii) and in the absence of Mg^{2+} (iv). Insets show structural cartoons of the state in which unfolding FECs start. Dashed lines are the WLC fits. Scale bar, 50 nm.

must have formed subsequently at too low a force to observe reliably. For the FECs without extreme-force unfolding events, the unfolding for most was consistent with starting in state I_1 containing native secondary structure only (Fig. 1d(ii), inset) and proceeding from I_3 , I_2 to I_4 to the unfolded state (U); the unfolding of the others was consistent with starting in state I_0 incorporating a nonnative helix P0 (Fig. 1d(iii), inset) and then proceeding through the same sequence of intermediates to U. Finally, the unfolding of the remainder matched the lengths expected for a nonnative state with all helices formed and the ring closed by PK but no threading of the 5' end through the ring (Fig. 1d(i), inset): first P1 was unfolded, then P2, P4 and finally PK. The refolding curves preceding these unfolding FECs were similar to those seen in the other cases (Supplementary Fig. 3), forming in the sequence I_4 , I_2 , I_3 and finally a variant of I_1 (denoted I_1^{PK}) with the PK intact but 5' end unthreaded.

The picture revealed by these measurements is intuitive: refolding involves forming each helix, in the order P4, P2, P3 and P1, followed by one of four options: (1) threading of the 5' end before PK closure to achieve the full native tertiary structure (N, 80% likelihood); (2) closure of PK before 5'-end threading to achieve only partial tertiary structure (I_1^{PK} , 4% likelihood); (3) neither threading nor PK closure, leaving an unconstrained 5' end (I_1 , 15% likelihood) or (4) alternative secondary-structure formation inhibiting 5'-end threading and PK formation (I_0 , 1% likelihood). The difference in threading between N and I_1^{PK} makes it impossible for them to interconvert without unfolding PK. The proportion of threaded and unthreaded states did not change noticeably when the waiting time between successive pulls was reduced from 5.5 to 0.5 s (Supplementary Fig. 4), indicating that threading and PK closure occurred rapidly.

To confirm this interpretation, FECs were remeasured in the presence of antisense oligonucleotides complementary to specific regions of the xrRNA, to prevent formation of the targeted substructures and interactions (Fig. 2a). Blocking P4 from forming with oligo 1 abolished the rip at roughly 20 pN attributed to P4 unfolding but left the resistant intermediate I_r intact, as expected (Fig. 2b(i)). In fact, the fraction of curves containing I_r increased slightly to roughly 90% using oligo 1, suggesting that steric hindrance from P4 may hinder 5'-end threading under normal conditions; the remaining approximately 10% of FECs started in a state with P1–P3 but no

PK (Extended Data Fig. 1). Blocking both P4 and PK with oligo 2, in contrast, abolished I_r entirely (Fig. 2b(ii)). FECs instead unfolded from state I_{5b} (P1–P3 and 5'-TC but no PK), at roughly 25–55 pN, or from state I_{1b} (P1–P3 only), at roughly 20 pN (Extended Data Fig. 2), indicating that the PK is essential for extreme mechanical resistance. These results show that 5'-TC can form independently of PK, and that the intermediate I_5 (Supplementary Fig. 2a) containing P1–P4 and 5'-TC but not PK—necessary for threading the 5' end before loop closure but not observed directly in the absence of antisense oligos—is indeed metastable. Last, blocking P1 and 5'-TC with oligo 3 abolished all tertiary structure (Fig. 2b(iii) and Extended Data Fig. 3), indicating that native PK formation requires P1. Remeasuring FECs without oligos but now removing the Mg^{2+} ions required for native 5'-TC³, most FECs (roughly 67%) revealed folding of only native secondary structure (Extended Data Fig. 4), similar to the results with oligo 3 (Extended Data Fig. 3). The remainder showed unfolding lengths and forces consistent with a variant PK lacking 5'-end threading, denoted PK' (Fig. 2b(iv) and Extended Data Fig. 4), suggesting PKs can form independently of 5'-TC. The length changes associated with the structural transitions in all FECs with oligo or without Mg^{2+} are listed in Supplementary Table 3.

Identifying key intermediates for mechanical resistance.

Combining these results provides a comprehensive view of the unfolding and refolding pathways (Fig. 3) and suggests the roles played by different parts of the xrRNA. The critical state responsible for the unprecedented mechanical stability is the intermediate I_r , not the full native state: P4 can unfold without compromising the mechanical strength of the ring knot. Mutational studies suggest that P4 is not essential to XR activity, either^{2,3}. Instead, P4 seems to play an assisting role as a 'buckle', helping to keep the ring compact and prepare the way for native 5'-TC and PK interactions by preventing the formation of the nonnative PKs such as PK' seen when P4 does not fold, which do not lead to threading (Extended Data Fig. 4). The other crucial state is I_1 , where the folding pathway partitions based on whether PK or 5'-TC forms first, since each can form independently of the other. If PK forms before the 5' end threads into the cleft in the three-helix junction formed by P1–P3, then no threading is possible and the RNA is topologically trapped in a state with lower mechanical stability. Only if the Mg^{2+} -mediated 5'-end

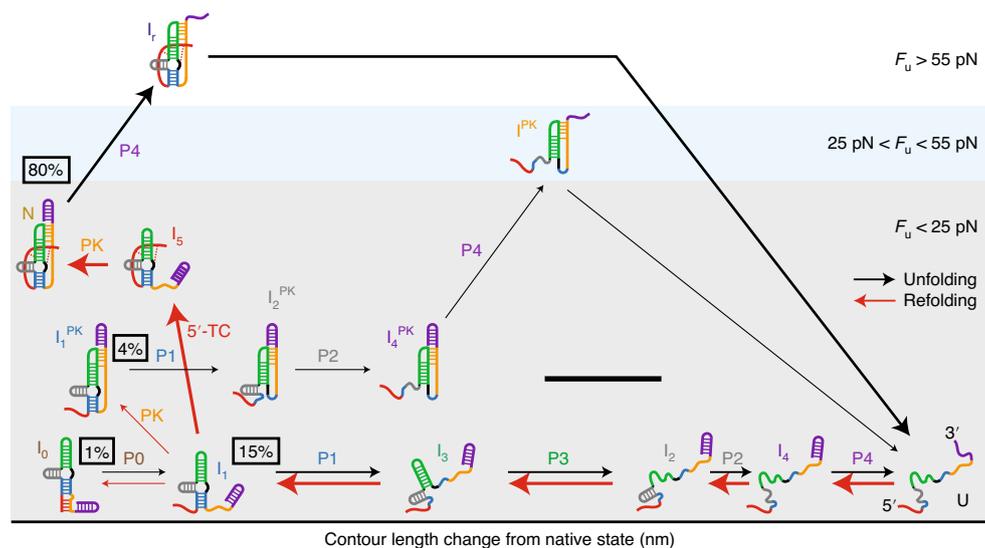


Fig. 3 | Unfolding and refolding pathways. The position of each state (shown as a cartoon of its structure) reflects its unfolding force (F_u) and ΔL_c relative to N. Arrows (black, unfolding; red, refolding) denote transitions, where the indicated structural element unfolds/refolds; their widths are proportional to the fraction of FECs in which the given transition occurs. Scale bar, 5 nm.

threading occurs before PK closure is the native state with exceptional mechanical resistance achieved.

Mutants with low-XR activity have low mechanical resistance.

Finally, to test if XR activity arises simply from the knot-like topology or if the extreme mechanical resistance of the fold is essential for resisting digestion, we studied the effects of two mutations: C22G, which reorganizes base pairing in P2–P3, and U4C, which disrupts 5'-TC stabilizing threading. Mutating C22 to G, which reduces XR activity from $80 \pm 4\%$ to $9 \pm 5\%$ (ref. ³), was proposed to replace the 5'-TC involving G3 and C44 with the base pair G22:C44, lengthening P3 by 1 base pair (bp) while disrupting the interactions stabilizing threading^{2,3}. However, only 32% of unfolding FECs were consistent with this scenario of a longer P3 and shorter P2: 15% containing secondary structure only (a variant of I_1) and 17% also featuring PK' (Fig. 4a, Extended Data Fig. 5). Most FECs, 68%, were consistent with the wild-type helix P3 and instead had P2 lengthened by a C9:G22 base pair (Fig. 4b, Extended Data Fig. 5 and Supplementary Table 4), leading to an increased unfolding force for P2 (Fig. 4c), whereas P1 was shortened by 1 bp; this alternate base pairing was confirmed from the lengths and forces observed when remeasuring the mutant in the presence of oligo 3 (Extended Data Fig. 6). Most of the curves with C9:G22 pairing (Fig. 4b, 53% of all FECs) again showed only secondary structure (another variant of I_1), but a substantial minority (15% of all FECs) retained a ring-knot architecture despite the distortions to P1–P2, and they displayed an extremely high-force intermediate after P4 unfolding analogous to I_r , denoted I_r^{m1} . I_r^{m1} can form here because C44 remains unpaired, allowing for full 5'-TC formation (Fig. 4b, inset and Extended Data Fig. 7). Hence this mutant is still capable of forming a ring knot with extreme mechanical resistance, but the mutation disfavors threading, greatly reducing the incidence of the resistant state.

Turning to the U4C mutant, disruption of the U4•A24:U42 triple is known to abolish XR in the Zika virus xrRNA²⁰, and similar mutations do the same for xrRNAs from other viruses;¹⁹ we confirmed that U4C reduced XR to 0% within our detection sensitivity (Extended Data Fig. 8a). Such mutations have been proposed to act by preventing the ring-knot structure from forming²⁰, however, it is unclear if they truly abolish the ring knot or if they merely weaken it mechanically. In fact, most FECs (52%) for the U4C mutant (Fig. 4d,

left and Supplementary Table 5) showed unfolding with a pattern of intermediates identical to that for unfolding the native wild-type ring knot, differing in only one respect: a much lower unfolding force, averaging only 38.0 ± 0.5 pN (Fig. 4e). This mutant can thus still form a ring knot by threading the 5' end into the helix junction and closing the PK around it (Extended Data Fig. 8), but the broken base triple greatly reduces the mechanical rigidity of the knot. While most of the remaining FECs (44%) unfolded from state I_1 , containing only native secondary structure (Fig. 4d, center), a small minority (1%) unfolded from the key intermediate I_5^{m2} , where the 5' end was threaded into the junction but PK remained open (Fig. 4d, right). This threaded but unknotted state was seen in most FECs when oligo 2 was added to prevent PK formation (Extended Data Fig. 8f), directly demonstrating that the U4C mutant still allows for 5'-end threading.

Finally, we compared the mechanical unfolding of the mutants and wild-type xrRNA to their XR activities. We found that XR (Fig. 4f, red) correlated well with the incidence of the extremely stable ring knot (Fig. 4f, black), for all three variants. However, it did not correlate with the incidence of all ring-knot folds when also including the ring knots lacking high mechanical stability (as seen for the U4C mutant) in the analysis (Fig. 4f, cyan). Extreme mechanical stability—not just the ring-knot fold topology—therefore appears to be a key ingredient needed to generate XR.

Discussion

These results provide insight into the central question of how the Zika virus xrRNA resists RNase digestion. The knot-like fold of this RNA was proposed to provide a structural basis for XR activity³: with the 5' end of the RNA protruding through a closed ring (Fig. 1a), host RNases that digest from the 5' end would be blocked from continuing once they reached the ring unless they could break it open to unravel the knot. In the simplest version of this picture, the knot-like topology is the essential element generating XR. Indeed, a recent study concluded from small-angle X-ray scattering measurements of mutants lacking XR activity, which appeared not to form ring knots, that the presence of a ring knot alone was sufficient to account for XR²⁰. However, by monitoring folding in the single-molecule regime, our measurements reveal a more nuanced picture: low-XR mutants do, in fact, form ring knots, but only some

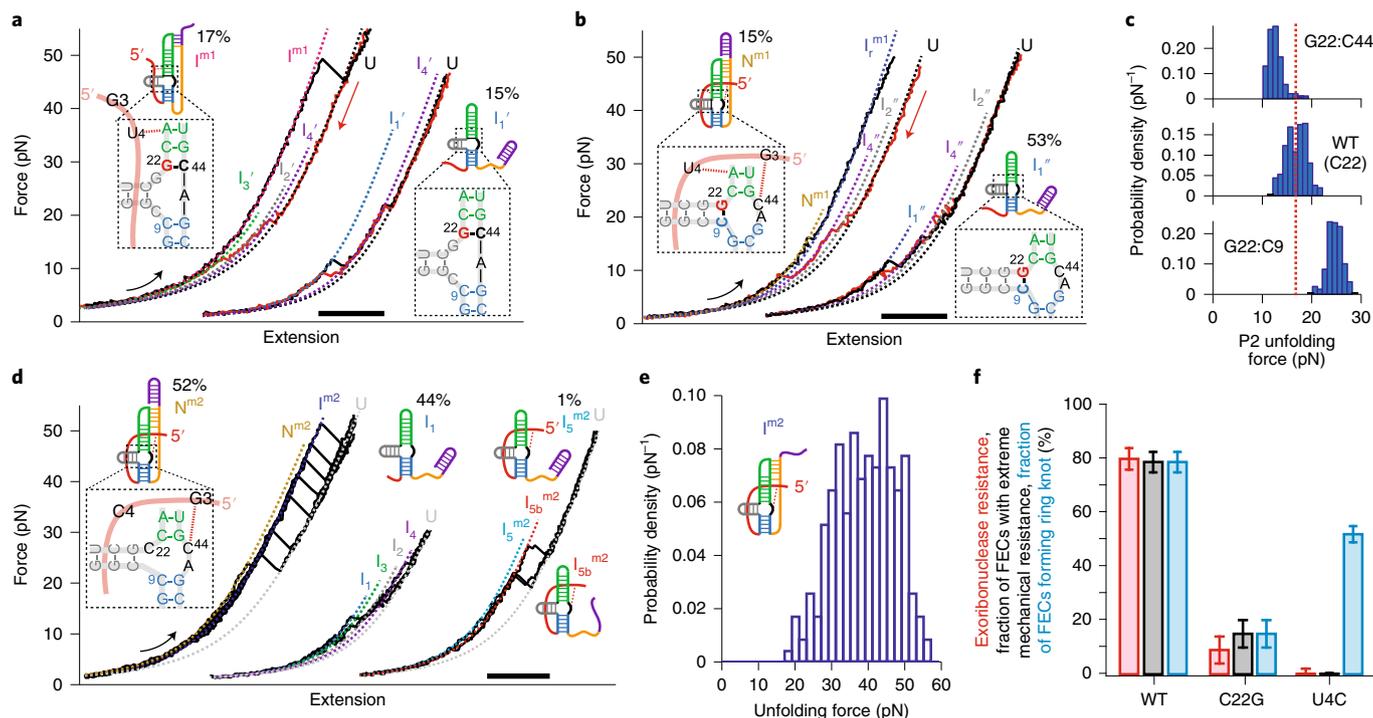


Fig. 4 | Folding of mutants with low XR. a–c, G22C mutant displays two different base pairings, one of which leads to the mechanically resistant knot a minority of the time. **a,** One-third of unfolding (black) and refolding (red) FECs were consistent with G22:C44 base pairing, which prevents 5'-end threading: 17% (left) featured the nonnative PK', 15% (right) featured secondary structure only. Insets show cartoons of structures at start of unfolding FECs. **b,** The remaining FECs were consistent with G22:C9 base pairing: most (53%, right) featured secondary structure only, but 15% (left) revealed a fully threaded ring knot. Insets show cartoons of structures at start of unfolding FECs. **c,** With G22:C44 base pairing (top, $N=105$), P2 is shorter than in wild-type (WT) (middle, $N=416$) and has a lower unfolding force, whereas with G22:C9 (bottom, $N=370$), P2 is longer and has a higher unfolding force. Dashed line, average unfolding force of wild-type P2. **d,e,** U4C mutant forms a ring knot with lower mechanical resistance. **d,** Most FECs (left) show unfolding of the full ring-knot structure but at much lower force than in the wild-type xrRNA. The remainder show unfolding of helices only (center), or more rarely the helices plus 5'-TC but no PK (right). **e,** The ring knot formed by the U4C mutant unfolds at an average force of 38 pN ($N=234$). **f,** XR (red) correlates strongly with occurrence of extreme mechanical resistance (state I_r , black) across all three variants, but not with ring-knot occurrence (cyan). XR values for WT and C22G from ref. ³, for U4C from an average of three replicates. Error bars for occurrences of ring-knot and extreme mechanical resistance represent standard error of binomial proportions, error bars for XR values of WT and C22G represent error reported in ref. ³, error bars for U4C represent detection sensitivity of assay. Scale bars in **a,b,d**, 50 nm.

of the time, and within a heterogeneous mixture of conformers. With U4C, the ring knot forms half the time, but it is mechanically much weaker than in the wild-type because of the disruptions to 5'-TC—indeed, its mechanical resistance is similar to that of simple RNA pseudoknots²⁶. The ring-knot fold thus does not by itself confer XR, else U4C would have roughly 50% XR activity, rather than none. Instead, it appears that extreme mechanical resistance is an essential requirement, reflected in the direct correspondence between the low but nonzero XR activity of C22G and the rate at which C22G forms a mechanically resistant ring knot, whose incidence is much reduced from the wild-type owing to altered base pairing in P2–P3.

The ring knot may thus enable the possibility for XR activity but is insufficient on its own to ensure it, whereas extreme mechanical resistance appears to be essential. Moderate mechanical resistance in the 40–50-pN range, such as provided by the weakened ring knot in U4C or the variant PK in C22G, is insufficient to confer RNase resistance, otherwise these mutants would exhibit much greater XR than they do. Instead, the extreme mechanical resistance of the native ring knot is needed. Reinforcing the importance of mechanical rather than thermodynamic stability, previous work showed that XR is not related simply to thermal stability, with some mutants lacking XR actually melting at higher temperatures than wild-type xrRNA²⁰.

Considering more explicitly the elements that determine sensitivity to RNase digestion, transient structural fluctuations exposing the 5' end are unlikely to be involved. Since the ring closed by PK is too small to allow passage of the RNase, digestion can only proceed if the ring is broken by opening PK. However, PK is itself stable, unfolding at forces of roughly 30–50 pN, and is thus very unlikely to break open spontaneously. Instead, the resistance to digestion appears to be connected intimately to the folding mechanism: it depends on how the RNA partitions³² between competing folding pathways leading to states that either have high mechanical resistance (and hence resist digestion) or do not. The wild-type xrRNA, for example, partitions between threaded and unthreaded conformers at state I_1 (Fig. 3); once PK closes the ring, the threaded and unthreaded conformers cannot interconvert and any unthreaded molecules are presumably then digested whereas those that are threaded to form the resistant ring knot remain undigested.

Turning to the factors that contribute to the mechanical stability required for XR, the U4C results indicate a critical role for 5'-TC, since disrupting 5'-TC lowers the unfolding force greatly while reducing ring-knot formation only modestly. Given that these contacts rigidify the 5' end by anchoring it to the three-helix junction (Fig. 1a), we speculate that disrupting 5'-TC allows the RNase to distort the ring more easily when it pulls on the 5' end, which eventually breaks open the PK holding the ring closed, deprotecting

the RNA from digestion and abolishing XR. Notably, unfolding by the RNase differs from how the tweezers unfold the knot: the tweezers apply force directly to the 3' end in addition to the 5' end, biasing the ring knot to unfold by opening PK, whereas the RNase pulls only on the 5' end while pushing against the ring, which provides worse mechanical leverage against PK. The true mechanical resistance of the native ring knot to unfolding by RNases is therefore almost certainly underestimated by the tweezers, as predicted computationally³³.

Although the correlation between XR and the extremely rigid ring knot is strong, a broader survey of mutants with different XR levels and xrRNAs from different viruses will be needed to establish more conclusively the central role of mechanical resistance suggested by this work and determine whether it is a general feature of all xrRNAs. Flaviviruses, for example, often feature two xrRNAs in tandem that are proposed to have different biological roles, with the upstream xrRNA being more resistant to digestion and acting as the primary roadblock to digestion, whereas the downstream xrRNA has lower resistance and is hypothesized to act as a modulator enhancing adaptation to the host^{34–36}. It is unclear if these different functional roles are reflected in different mechanisms for generating XR (including different folding mechanisms). Another important question is whether XR requires an unfolding force above some threshold, for example the RNase stall force. Since Xrn1 can unfold and digest the U4C ring knot, its stall force is presumably over roughly 50 pN (Fig. 4e), higher than the 15–25 pN stall forces of other processive RNA enzymes such as RNA polymerase³⁷ and the ribosome³⁸. However, stall forces have not yet been reported for RNases.

Whereas many proteins act primarily as mechanical elements in the cell, likewise this xrRNA appears to be an example of an RNA that does so too. Its action as a roadblock resisting RNase digestion mechanically adds yet another entry to the growing repertoire of capabilities displayed by RNA. Comparing this knot-like RNA to knotted proteins, a key difference is that the ring knot helps to generate the unusual mechanical stability of the xrRNA, which is central to its function; in contrast, knots are typically only incidental to function in proteins^{39–41}. Unlike in proteins, the xrRNA ring knot is not a true topological knot, where the backbone cannot be straightened without breaking it; the ring knot can be unraveled by opening PK to unthread the 5' end. Indeed, topological knots should not form in RNA, as they would prevent the operation of essential processive enzymes. It remains to be seen if other folds lacking knot-like topology can also generate extreme mechanical resistance and high levels of XR activity. If so, based on the results above we might expect them to feature dense networks of tertiary contacts rigidifying the 5' end and sequestering it from RNases.

Finally, the xrRNA folding mechanism reveals that an intricate pattern of intermediate states must be formed sequentially to arrive at the ring knot, lest the threading of the 5' end go awry and the xrRNA instead fold into a nonresistant structure. This result suggests that ring-knot formation could be targeted therapeutically. For example, considering the intermediate I_1 that mediates threading, a ligand that binds in the cleft of the three-helix junction could prevent 5'-end threading, especially if it interacts with the bases forming 5'-TC, thereby redirecting the folding away from the resistant state. The complexity of the full folding pathway holds out the hope that there might be several stages that are susceptible to intervention in this way. Such a therapeutic strategy may also be applicable to the many other important viruses featuring xrRNAs, if intermediates regulating 5'-end threading are a shared hallmark of the folding.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of

author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41589-021-00829-z>.

Received: 27 July 2020; Accepted: 3 June 2021;

References

- Steckelberg, A.-L. et al. A folded viral noncoding RNA blocks host cell exoribonucleases through a conformationally dynamic RNA structure. *Proc. Natl Acad. Sci. USA* **115**, 6404–6409 (2018).
- Chapman, E. G. et al. The structural basis of pathogenic subgenomic flavivirus RNA (sRNA) production. *Science* **344**, 307–310 (2014).
- Akiyama, B. M. et al. Zika virus produces noncoding RNAs using a multi-pseudoknot structure that confounds a cellular exonuclease. *Science* **354**, 1148–1152 (2016).
- Steckelberg, A.-L., Vicens, Q., Costantino, D. A., Nix, J. C. & Kieft, J. S. The crystal structure of a Poliovirus exoribonuclease-resistant RNA shows how diverse sequences are integrated into a conserved fold. *RNA* **26**, 1767–1776 (2020).
- Pijlman, G. P. et al. A highly structured, nuclease-resistant, noncoding RNA produced by flaviviruses is required for pathogenicity. *Cell Host Microbe* **4**, 579–591 (2008).
- Funk, A. et al. RNA structures required for production of subgenomic flavivirus RNA. *J. Virol.* **84**, 11407–11417 (2010).
- Dilweg, I. W., Gulyaev, A. P. & Olsthoorn, R. C. Structural features of an Xrn1-resistant plant virus RNA. *RNA Biol.* **16**, 838–845 (2019).
- Gunawardene, C. D., Newburn, L. R. & White, K. A. A 212-nt long RNA structure in the tobacco necrosis virus-D RNA genome is resistant to Xrn degradation. *Nucleic Acids Res.* **47**, 9329–9342 (2019).
- Steckelberg, A.-L., Vicens, Q. & Kieft, J. S. Exoribonuclease-resistant RNAs exist within both coding and noncoding subgenomic RNAs. *mBio* **9**, e02461 (2018).
- Jinek, M., Coyle, S. M. & Doudna, J. A. Coupled 5' nucleotide recognition and processivity in Xrn1-mediated mRNA decay. *Mol. Cell* **41**, 600–608 (2011).
- Nagarajan, V. K., Jones, C. I., Newbury, S. F. & Green, P. J. XRN 5'→3' exoribonucleases: structure, mechanisms and functions. *Biochim. Biophys. Acta* **1829**, 590–603 (2013).
- Bavia, L., Mosimann, A. L. P., Aoki, M. N. & Duarte dos Santos, C. N. A glance at subgenomic flavivirus RNAs and microRNAs in flavivirus infections. *Virol. J.* **13**, 84 (2016).
- Moon, S. L. et al. A noncoding RNA produced by arthropod-borne flaviviruses inhibits the cellular exoribonuclease XRN1 and alters host mRNA stability. *RNA* **18**, 2029–2040 (2012).
- Schnettler, E. et al. Noncoding flavivirus RNA displays RNA interference suppressor activity in insect and mammalian cells. *J. Virol.* **86**, 13486–13500 (2012).
- Manokaran, G. et al. Dengue subgenomic RNA binds TRIM25 to inhibit interferon expression for epidemiological fitness. *Science* **350**, 217–221 (2015).
- Pompon, J. et al. Dengue subgenomic flaviviral RNA disrupts immunity in mosquito salivary glands to increase virus transmission. *PLoS Path.* **13**, e1006535 (2017).
- Göertz, G. P. et al. Subgenomic flavivirus RNA binds the mosquito DEAD/H-box helicase ME31B and determines Zika virus transmission by *Aedes aegypti*. *Proc. Natl Acad. Sci. USA* **116**, 19136–19144 (2019).
- Gould, E. A. & Solomon, T. Pathogenic flaviviruses. *Lancet* **371**, 500–509 (2008).
- MacFadden, A. et al. Mechanism and structural diversity of exoribonuclease-resistant RNA structures in flaviviral RNAs. *Nat. Commun.* **9**, 119 (2018).
- Niu, X. et al. Molecular mechanisms underlying the extreme mechanical anisotropy of the flaviviral exoribonuclease-resistant RNAs (xrRNAs). *Nat. Commun.* **11**, 5496 (2020).
- Ritchie, D. B. & Woodside, M. T. Probing the structural dynamics of proteins and nucleic acids with optical tweezers. *Curr. Opin. Struct. Biol.* **34**, 43–51 (2015).
- Wang, M. D., Yin, H., Landick, R., Gelles, J. & Block, S. M. Stretching DNA with optical tweezers. *Biophys. J.* **72**, 1335–1346 (1997).
- Onoa, B. et al. Identifying kinetic barriers to mechanical unfolding of the *T. thermophila* ribozyme. *Science* **299**, 1892–1895 (2003).
- Li, P. T. X., Bustamante, C. & Tinoco, I. Real-time control of the energy landscape by force directs the folding of RNA molecules. *Proc. Natl Acad. Sci. USA* **104**, 7039–7044 (2007).
- Greenleaf, W. J., Frieda, K. L., Foster, D. A. N., Woodside, M. T. & Block, S. M. Direct observation of hierarchical folding in single riboswitch aptamers. *Science* **319**, 630–633 (2008).

26. Ritchie, D. B., Foster, D. A. N. & Woodside, M. T. Programmed -1 frameshifting efficiency correlates with RNA pseudoknot conformational plasticity, not resistance to mechanical unfolding. *Proc. Natl Acad. Sci. USA* **109**, 16167–16172 (2012).
27. Halma, M. T. J., Ritchie, D. B., Cappellano, T. R., Neupane, K. & Woodside, M. T. Complex dynamics under tension in a high-efficiency frameshift stimulatory structure. *Proc. Natl Acad. Sci. USA* **116**, 19500–19505 (2019).
28. Anthony, P. C., Perez, C. F., García-García, C. & Block, S. M. Folding energy landscape of the thiamine pyrophosphate riboswitch aptamer. *Proc. Natl Acad. Sci. USA* **109**, 1485–1489 (2012).
29. Chen, G., Chang, K.-Y., Chou, M.-Y., Bustamante, C. & Tinoco, I. Triplex structures in an RNA pseudoknot enhance mechanical stability and increase efficiency of -1 ribosomal frameshifting. *Proc. Natl Acad. Sci. USA* **106**, 12706–12711 (2009).
30. Smith, S. B., Cui, Y. & Bustamante, C. Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules. *Science* **271**, 795–799 (1996).
31. Woodside, M. T. et al. Nanomechanical measurements of the sequence-dependent folding landscapes of single nucleic acid hairpins. *Proc. Natl Acad. Sci. USA* **103**, 6190–6195 (2006).
32. Thirumalai, D., Lee, N., Woodson, S. A. & Klimov, D. Early events in RNA folding. *Annu. Rev. Phys. Chem.* **52**, 751–762 (2001).
33. Suma, A., Coronel, L., Bussi, G. & Micheletti, C. Directional translocation resistance of Zika xrRNA. *Nat. Commun.* **11**, 3749 (2020).
34. Kieft, J. S., Rabe, J. L. & Chapman, E. G. New hypotheses derived from the structure of a flaviviral Xrn1-resistant RNA: Conservation, folding, and host adaptation. *RNA Biol.* **12**, 1169–1177 (2015).
35. Villordo, S. M., Filomatori, C. V., Sánchez-Vargas, I., Blair, C. D. & Gamarnik, A. V. Dengue virus RNA structure specialization facilitates host adaptation. *PLoS Path.* **11**, e1004604 (2015).
36. Chapman, E. G., Moon, S. L., Wilusz, J. & Kieft, J. S. RNA structures that resist degradation by Xrn1 produce a pathogenic Dengue virus RNA. *eLife* **3**, e01892 (2014).
37. Yin, H. et al. Transcription against an applied force. *Science* **270**, 1653–1657 (1995).
38. Liu, T. et al. Direct measurement of the mechanical work during translocation by the ribosome. *eLife* **3**, e03406 (2014).
39. Ziegler, F. et al. Knotting and unknotting of a protein in single molecule experiments. *Proc. Natl Acad. Sci. USA* **113**, 7533–7538 (2016).
40. Bustamante, A. et al. The energy cost of polypeptide knot formation and its folding consequences. *Nat. Commun.* **8**, 1581 (2017).
41. San Martín, Á. et al. Knots can impair protein degradation by ATP-dependent proteases. *Proc. Natl Acad. Sci. USA* **114**, 9864–9869 (2017).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2021

Methods

Sample preparation. Constructs for SMFS measurements were generated following methods described previously²⁶. Briefly, the sequence for the wild-type Zika virus xrRNA (or C22G and U4C mutants) was inserted into a pcDNA5/FRT/TO plasmid between the HindIII and BamHI restriction sites (Supplementary Table 1). The resulting transcription template, which contained the xrRNA flanked by linker regions on either side, was then amplified by PCR and transcribed *in vitro* using T7 RNA polymerase. Two ssDNA handles (one labeled with biotin and complementary to the 840 nt on the 3' end of the transcript, the other labeled with digoxigenin and complementary to the 2,107 nt on the 5' end of the transcript) were produced by asymmetric PCR from double-stranded DNA PCR products corresponding to the flanking handle sequences. The handles were thermally annealed with the RNA transcript, then incubated with 600- and 820-nm diameter polystyrene beads labeled with avidin DN and antidigoxigenin, respectively, to create dumbbells. Dumbbells were placed in measuring buffer (50 mM MOPS pH 7.0, 130 mM KCl, 4 mM MgCl₂, 50 U ml⁻¹ Suprase•In RNase inhibitor (Ambion)) containing oxygen-scavenging system (40 U ml⁻¹ glucose oxidase, 185 U ml⁻¹ catalase and 8.3 mg ml⁻¹ glucose) and inserted into a sealed sample chamber on a clean microscope slide in the optical trap. These ionic conditions are near-physiological⁴², with Mg²⁺ concentration in the range where changes have little effect on folding¹. For measurements without Mg²⁺, the MgCl₂ was removed from the measuring buffer and EDTA added in its stead, to a final concentration of 1 mM. For measurements using antisense oligos, the desired oligo was added to the measuring buffer at a final concentration of 10 μM.

SMFS measurements and analysis. FECs were measured with a custom-built, dual-beam optical trap similar to one described previously⁴³, controlled by Labview 2018. Briefly, two orthogonally polarized beams from a single 1,064-nm laser were steered independently with acousto-optic deflectors to create two traps. The motions of beads held in the traps were detected by collecting, from two orthogonally polarized 830-nm laser beams aligned on the traps, the light that was scattered by the beads onto position-sensitive diodes. The trap stiffnesses (0.54 and 0.56 pN nm⁻¹), were calibrated as described⁴⁴. Data were sampled at 20 kHz and filtered online at the Nyquist frequency with an eight-pole Bessel filter.

Contour length changes were determined from the FECs by fitting each state in the FECs to an extensible WLC model consisting of two WLCs in series: one for the duplex handles, and one for the amount of single-stranded RNA unfolded in each state. The extensible WLC is given by

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

where L_p is the persistence length, L_c the contour length and K the elastic modulus²². The parameters L_p , L_c and K describing the duplex handles were first determined by fitting the FEC for the natively folded state (N). Then, the FECs for the various intermediate states and unfolded states were fit by treating L_p , L_c and K as fixed variables for both the duplex handles (values determined from the previous fit) and the unfolded ssRNA (values determined from the literature: $L_p = 1$ nm, $K = 2,000$ pN)⁴⁵. The contour length of unfolded ssRNA was therefore the only free parameter for fitting the intermediate and unfolded states. For those FECs that did not contain N , the unfolded state was instead used as the reference point for WLC fitting to measure the contour length of RNA unfolded in partially folded intermediates (Supplementary Tables 2–4). To decrease the uncertainty in the fitting parameters, we did not fit FECs individually and average the results, but rather combined all the FECs showing the same behavior from a given set of measurements made using the same trap calibration parameters on a given molecule, and then fit each branch of the combined FECs to equation (1). The results were then averaged across the different sets of combined FECs. All FEC analysis was done using Igor Pro v.7.08 and Origin v.8.5.

Structures were identified from FECs based primarily on the contour length changes (ΔL_c) found from WLC fits, which we compared to the values expected for the proposed structures. The total ΔL_c for completely unfolding a structure containing n_m nucleotides is given by

$$\Delta L_c = n_m \times L_c^{nt} - d_T \quad (2)$$

where $L_c^{nt} = 0.59$ nm nt⁻¹ is the contour length per nt (ref. ⁴⁶) and d_T is the distance between the termini of the folded structure (as measured or estimated from the reported crystal structure). For example, natively folded Zika virus xrRNA has $d_T = 3.0$ nm and $n_m = 70$, leading to $\Delta L_c^{total} = 38.3$ nm for complete unfolding. Individual helices were assumed to have $d_T = 2.2$ nm, the width of an A-form helix⁴⁶.

Structural assignments were also informed by the unfolding and refolding forces observed. Unfolding and refolding of secondary structures alone (helices and hairpins), without any tertiary structures, typically occurs at forces in the range 10–25 pN, relatively close to equilibrium (so that unfolding and refolding forces are similar) and with a narrow force distribution, owing to the fact that duplexes are not very rigid with respect to unzipping³¹. In contrast, tertiary contacts increase the rigidity of RNA structures very noticeably, leading to higher unfolding forces (typically above 25 pN), a broader distribution of unfolding forces, and stronger hysteresis when refolding^{23–27}. Elements of secondary structure that do

not participate in tertiary contacts are expected to unfold independently, usually in the order of their thermodynamic stability^{25,31}, whereas those that form part of larger-scale tertiary structures may be protected by the tertiary contacts so that they only unfold at the higher forces at which the tertiary contacts are broken^{23,25,27,47}. Hence for the unthreaded native PK, which has tertiary contacts only with P3, P3 unfolds simultaneously with PK and after the other helices have unfolded, whereas for state I₁ (which lacks any tertiary contacts), P3 unfolds before P2 and P4 because of its lower thermodynamic stability. Note that only structural changes that alter the end-to-end length of the RNA can be observed directly by SMFS; conformational changes that leave the end-to-end length unchanged (for example, breathing of base pairs or junctions) will remain undetected unless they also change the unfolding force sufficiently so that distinct unfolding force distributions can be identified.

XR activity assays. The sensitivity to digestion by Xrn1 was measured as described previously²³. To facilitate loading of Xrn1, an overhang of ssRNA was added at the 5' end of the xrRNA sequences, creating 134-nt constructs (Supplementary Table 1). RNA transcripts were treated with pyrophosphohydrolase (RppH) and purified on MinElute columns to generate input RNA samples with 5'-monophosphate ends. Roughly 120 ng of input RNA was mixed with 1 μl of 10-μM spacer DNA oligo, complementary to part of the single-strand overhang (Supplementary Table 1), in 15.5 μl of reaction buffer (50 mM Tris pH 7.9, 100 mM NaCl, 10 mM MgCl₂, 1 mM DTT). The mixture was heated to 90 °C for 3 min before refolding the RNA at room temperature for 20 min, adding 0.5 μl of murine RNase inhibitor (40 U/μl, New England Biolabs) to prevent nonspecific degradation. Then, 2 μl of 1 U μl⁻¹ Xrn1 (New England Biolabs) or 2 μl of reaction buffer was added to 8 μl of the mixture containing RNA, respectively for (+) and (–) Xrn1 measurements. The reactions were incubated for 2 h at 37 °C, and terminated by adding 5 μl of formamide loading buffer and heating to 90 °C for 3 min, before loading the products into an 8% denaturing polyacrylamide gel. Gels were stained with SYBR Green II (Invitrogen) and quantified using ImageJ v.1.8.0 (ref. ⁴⁸).

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The data supporting the findings of this work have been deposited in Figshare (<https://doi.org/10.6084/m9.figshare.14544495>). Source data are provided with this paper.

References

- Romani, A. M. Magnesium homeostasis in mammalian cells. *Front Biosci.* **12**, 308–331 (2007).
- Neupane, K., Yu, H., Foster, D. A. N., Wang, F. & Woodside, M. T. Single-molecule force spectroscopy of the add adenine riboswitch relates folding to regulatory mechanism. *Nucleic Acids Res.* **39**, 7677–7687 (2011).
- Neuman, K. C. & Block, S. M. Optical trapping. *Rev. Sci. Instrum.* **75**, 2787–2809 (2004).
- Seol, Y., Skinner, G. M. & Visscher, K. Elastic properties of a single-stranded charged homopolymeric ribonucleotide. *Phys. Rev. Lett.* **93**, 118102 (2004).
- Saenger, W. *Principles of Nucleic Acid Structure* (Springer, 1984).
- Halma, M. T. J., Ritchie, D. B. & Woodside, M. T. Conformational Shannon entropy of mRNA structures from force spectroscopy measurements predicts the efficiency of -1 programmed ribosomal frameshift stimulation. *Phys. Rev. Lett.* **126**, 038102 (2021).
- Abramoff, M. D., Magalhaes, P. J. & Ram, S. J. Image processing with ImageJ. *Biophotonics Int.* **11**, 36–42 (2004).

Acknowledgements

We thank J. Kieft for helpful discussions about xrRNAs and D. Ritchie for assistance in designing the optical tweezers assay. This work was supported by the Canadian Institutes of Health Research grant reference number MOP-142449 (to M.T.W.), Alberta Innovates iCORE Strategic Chair (to M.T.W.) and the National Research Council Canada (to M.T.W.).

Author contributions

M.Z. and M.T.W. designed the research. M.Z. prepared samples and performed measurements. M.Z. and M.T.W. analyzed the data. M.Z. and M.T.W. wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

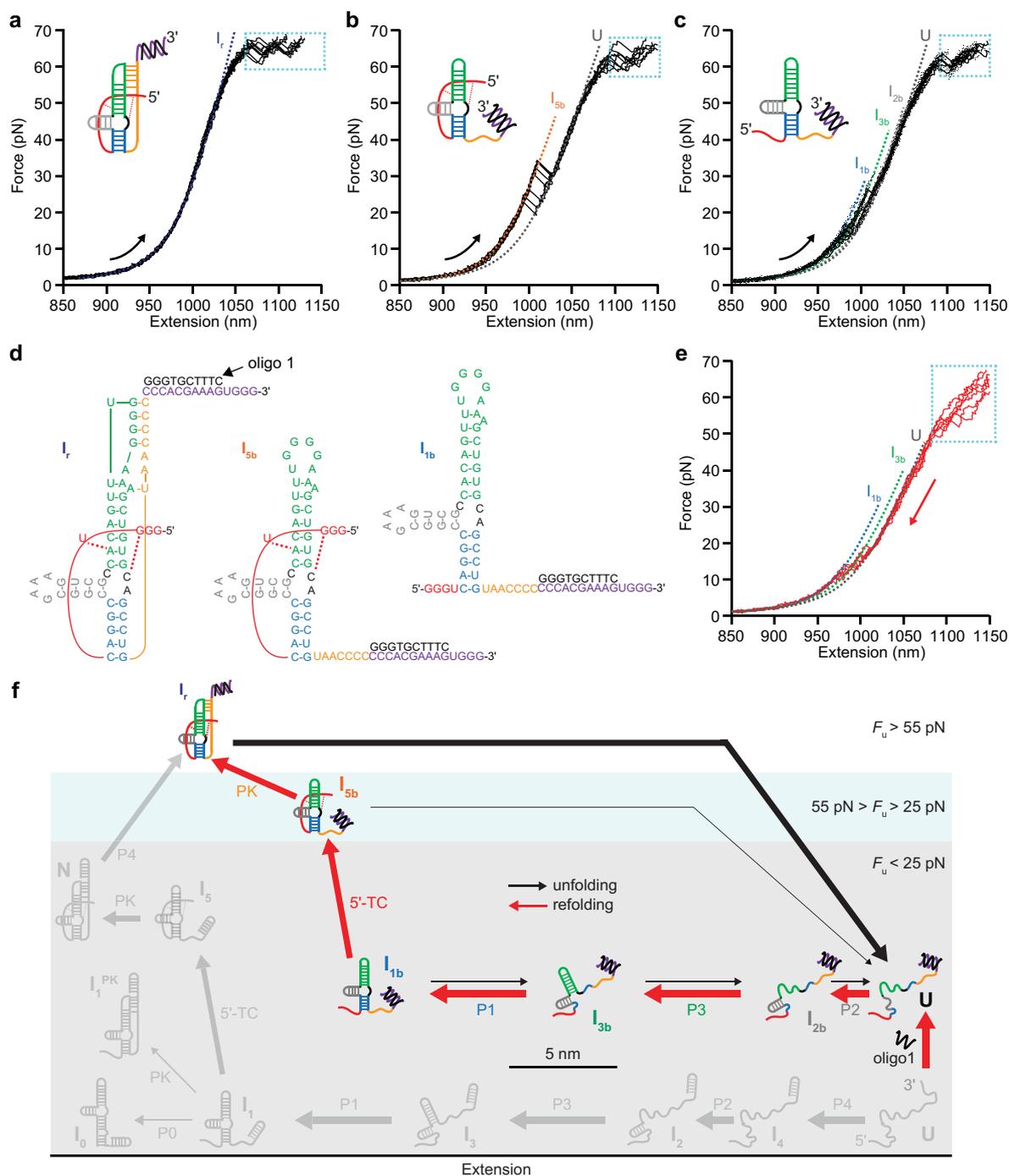
Extended data is available for this paper at <https://doi.org/10.1038/s41589-021-00829-z>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41589-021-00829-z>.

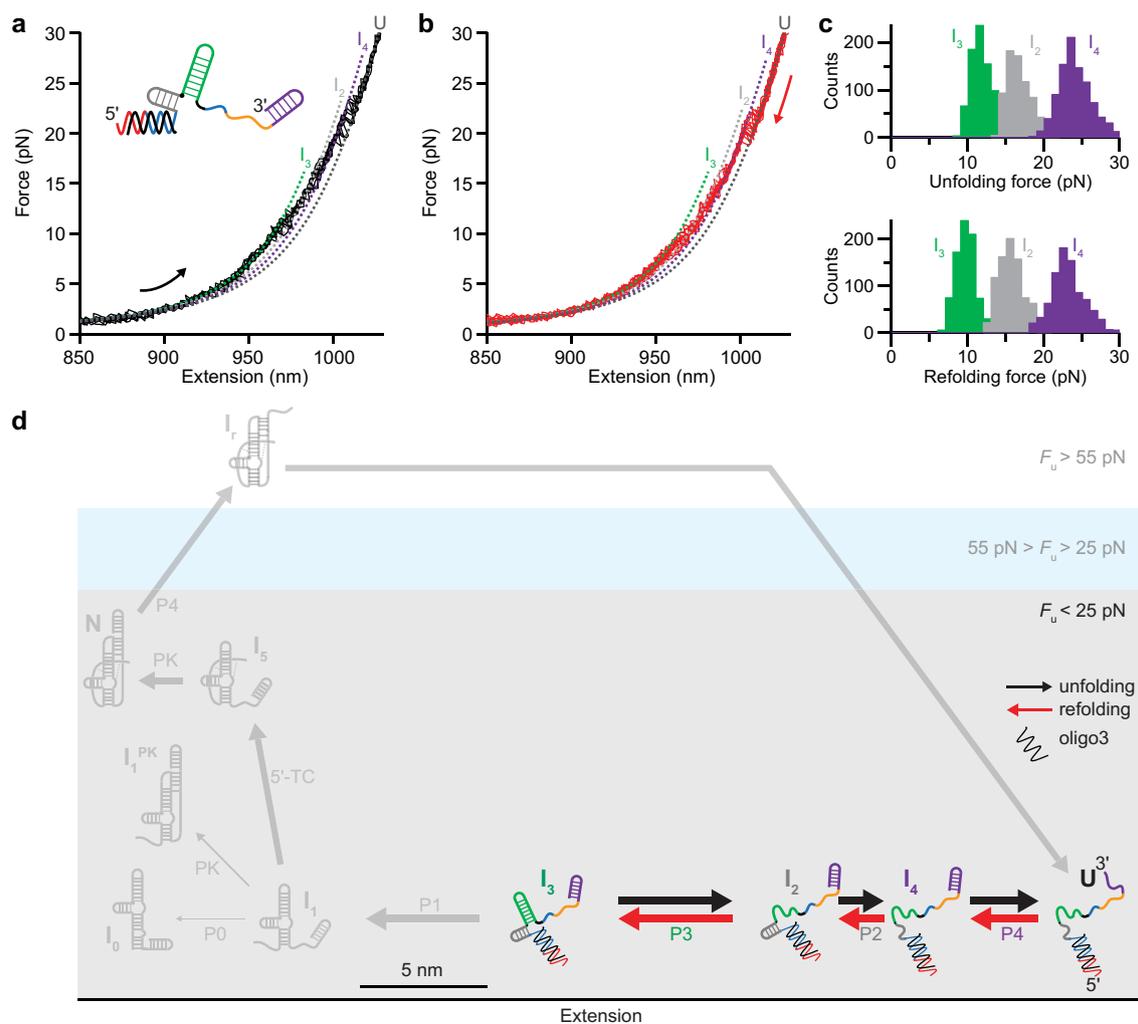
Correspondence and requests for materials should be addressed to M.T.W.

Peer review information *Nature Chemical Biology* thanks Pan T. X. Li and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

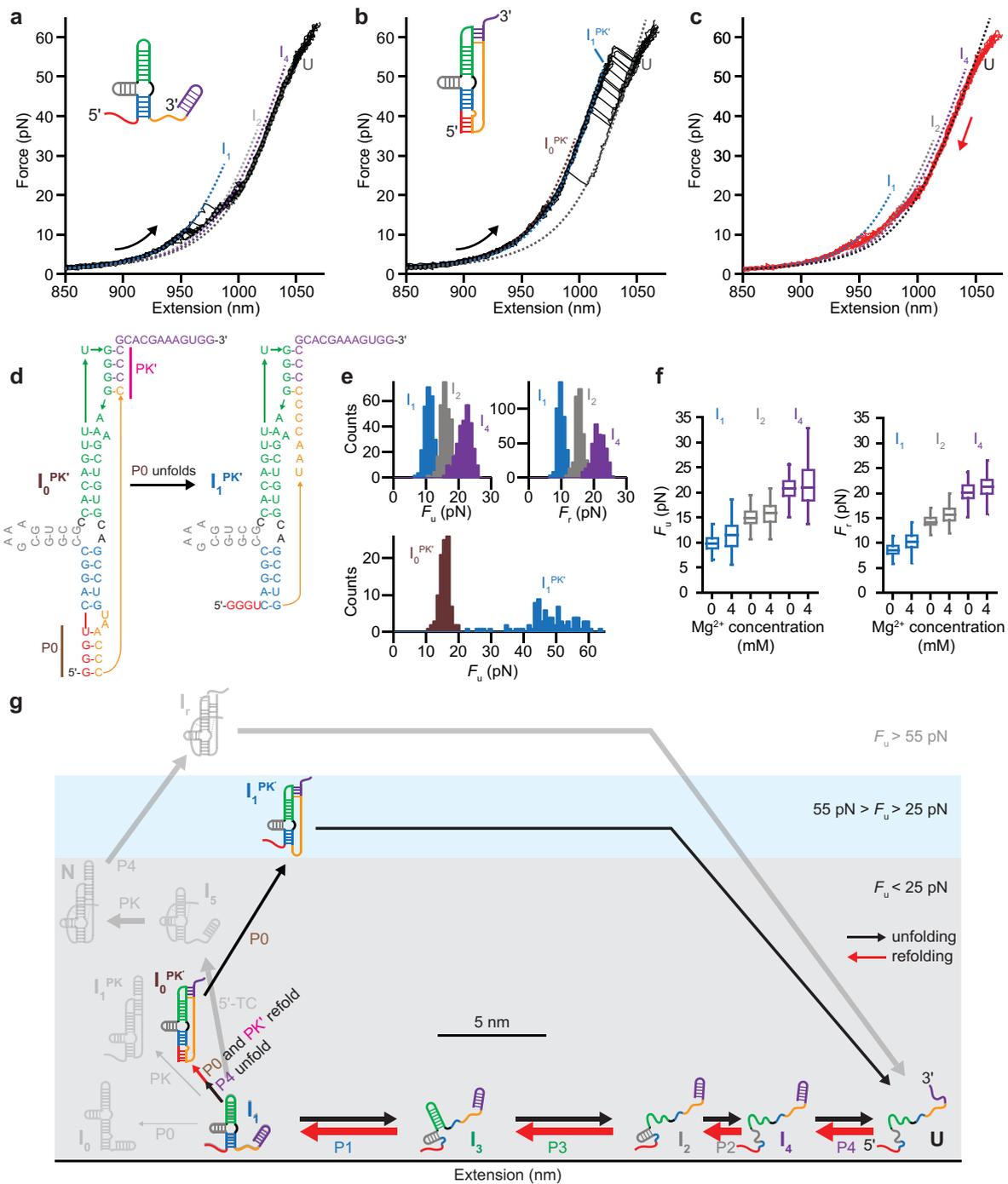
Reprints and permissions information is available at www.nature.com/reprints.



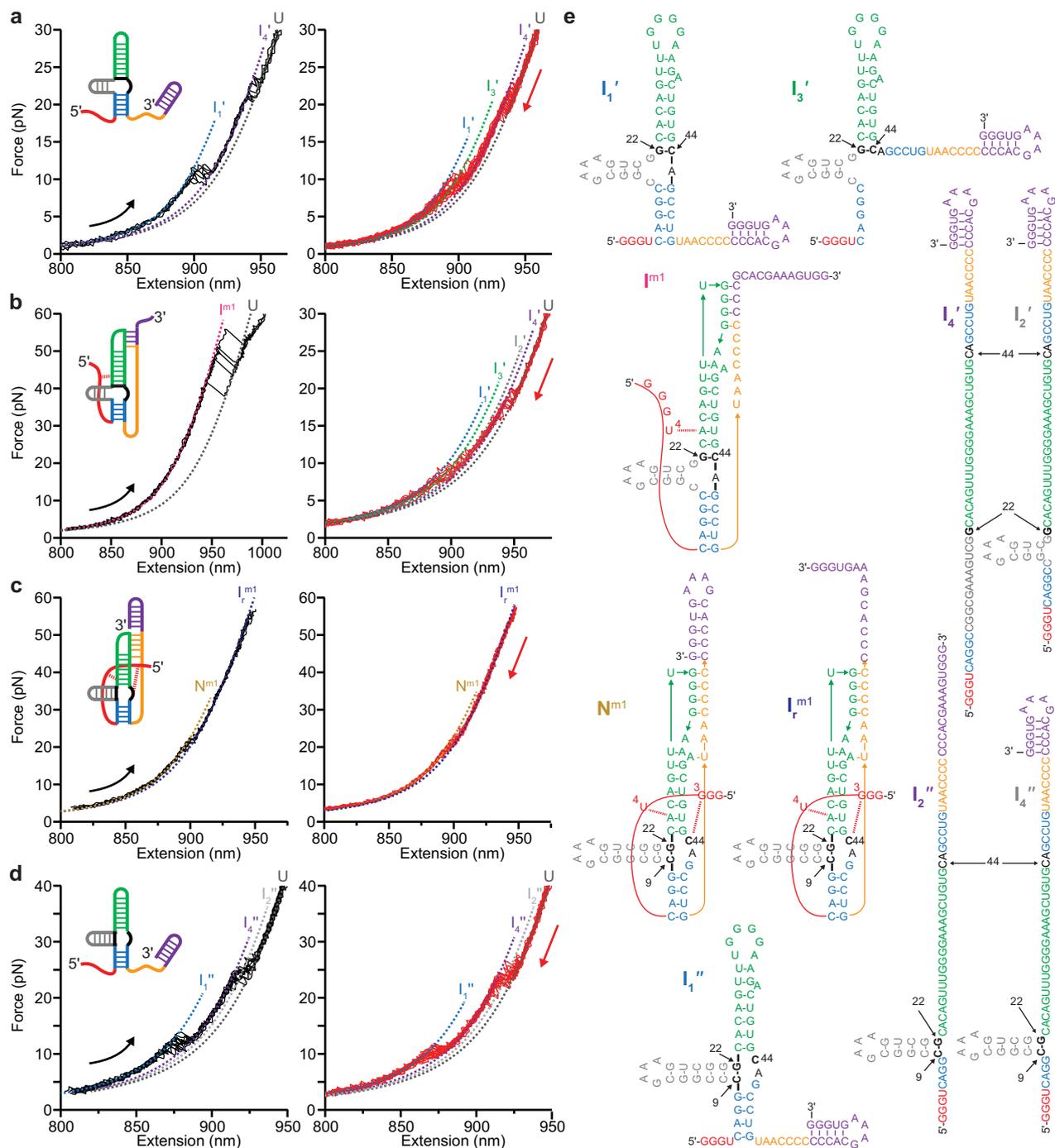
Extended Data Fig. 1 | Effects of anti-sense oligo 1 on SMFs of Zika virus xrRNA. a, Most unfolding FECs ($90 \pm 2\%$) show the extreme mechanical stability of I_r but no preceding ~ 20 -pN unfolding event, confirming that P4 unfolds prior to I_r . Dashed lines show WLC fits to each state, cyan boxes show overstretching region, inset shows cartoon of structure before unfolding. **b**, A very few FECs ($\sim 0.3\%$) show unfolding forces in the range ~ 25 - 50 pN indicative of tertiary structure, and ΔL_c consistent with state I_{sb} , where the 5' end is threaded but the ring is not closed. **c**, Some FECs ($10\% \pm 2\%$) show unfolding of secondary structure only, consistent with state I_{1b} . **d**, Proposed structures of I_r , I_{sb} and I_{1b} with oligo 1 bound. Contour length changes expected from unfolding each structure are listed in Supplementary Table 3. **e**, Most refolding FECs show the formation of I_{1b} ; in some cases, transitions into I_r and/or I_{sb} are seen directly (for example in Fig. 2b), in other cases, they occur at sufficiently low force to remain undetected. **f**, Unfolding (black) and refolding (red) pathways in the presence of oligo 1. Arrow thicknesses indicate pathway probabilities. For comparison, the states and transitions prevented by binding of oligo 1 are shown in grey.



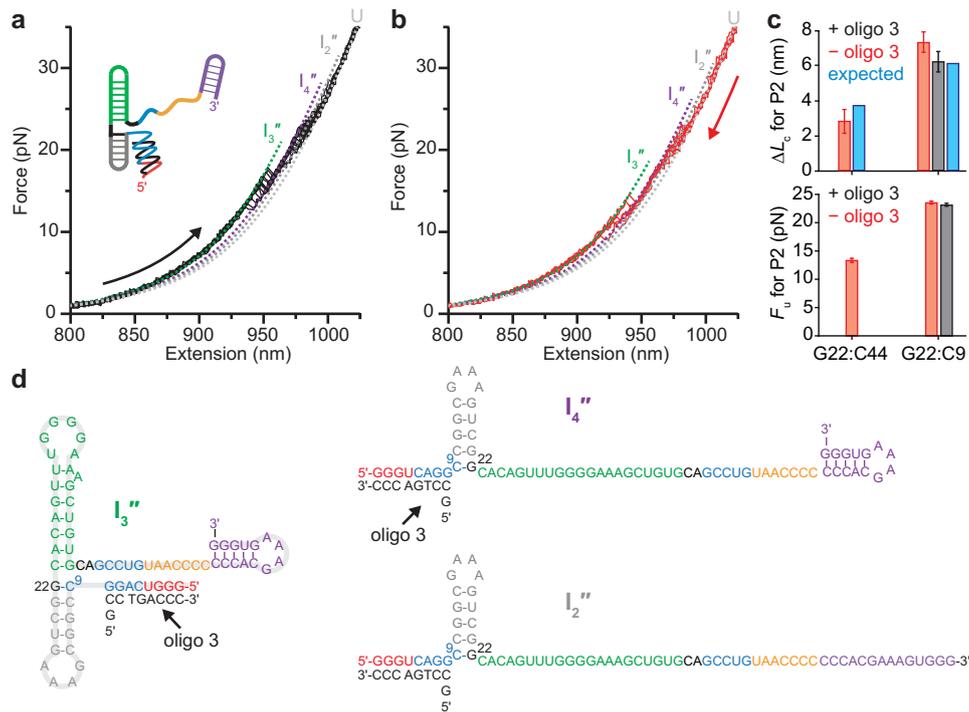
Extended Data Fig. 3 | Effects of anti-sense oligo 3 on SMFS of Zika virus xrRNA. **a, b**, All of the unfolding (black) and refolding (red) FECs show the low-force transitions characteristic of secondary structure only, with length changes indicative of state I_3 . Dashed lines show WLC fits to different states, inset shows structure before unfolding with oligo 3 bound. **c**, Unfolding (top panel) and refolding (bottom panel) force distributions show low hysteresis characteristic of helix unfolding/refolding. **d**, Unfolding (black) and refolding (red) pathways in the presence of oligo 3. Arrow thicknesses indicate pathway probabilities. For comparison, the states and transitions blocked by oligo 3 binding are shown in grey.



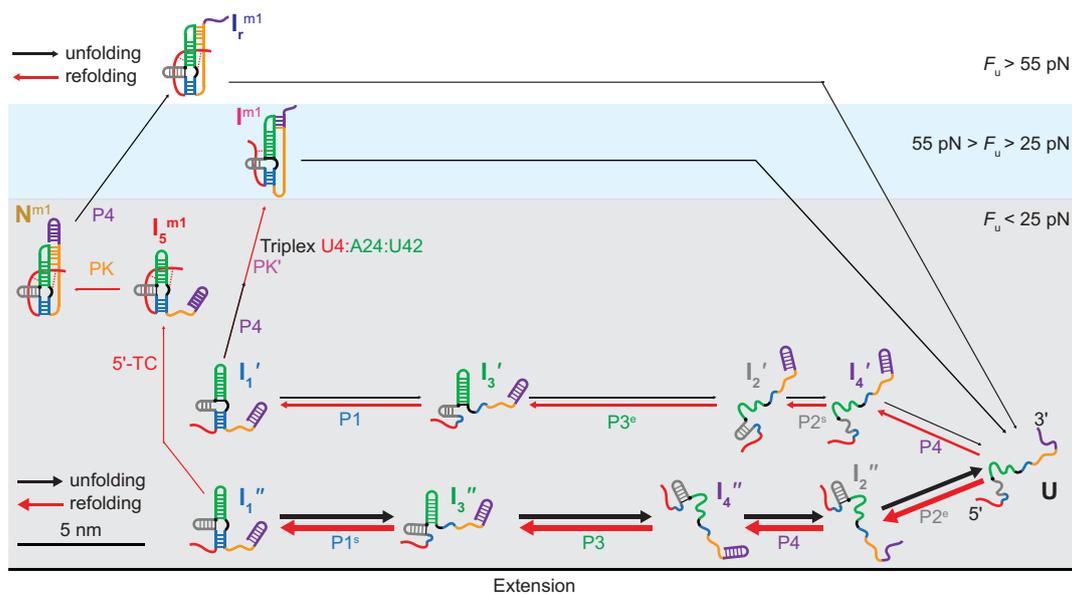
Extended Data Fig. 4 | SMFS of Zika virus xrRNA in the absence of Mg²⁺. **a**, Two-thirds of unfolding FECs (~67%) show low-force transitions consistent with the unfolding of I₁ (Supplementary Fig. 2b). Dashed lines show WLC fits to each state, inset shows structure before unfolding. **b**, One-third of unfolding FECs (~33%) show high-force transitions consistent with state I₀^{PK} containing a non-native pseudoknot that inhibits the folding of P4. **c**, All refolding FECs show sequential formation of I₄ → I₂ → I₁, indicating I₁ is Mg²⁺-independent and suggesting I₀^{PK} is derived from I₁. **d**, Cartoon of the I₀^{PK} → I₁^{PK} transition. Expected contour length changes listed in Supplementary Table 3. **e**, Unfolding (F_u) and refolding (F_r) force distributions of the states seen in FECs. High unfolding forces for I₁^{PK} indicate the presence of tertiary interactions. **f**, Unfolding and refolding forces for I₄, I₂ and I₁ show a small stabilization effect from Mg²⁺. Center line indicates median value, box edges indicate 25th and 75th percentiles, whiskers indicate 5th and 95th percentiles. N = 283, 287, and 285 respectively for unfolding I₁, I₂, and I₄ without Mg²⁺; N = 339, 328, and 352 for unfolding with Mg²⁺; N = 386, 386, 369 for refolding without Mg²⁺; and N = 477, 475, and 476 for refolding with Mg²⁺. **g**, Unfolding (black) and refolding (red) pathways in the absence of Mg²⁺. Arrow thicknesses indicate pathway probabilities. For comparison, the states and transitions dependent on Mg²⁺ are shown in grey.



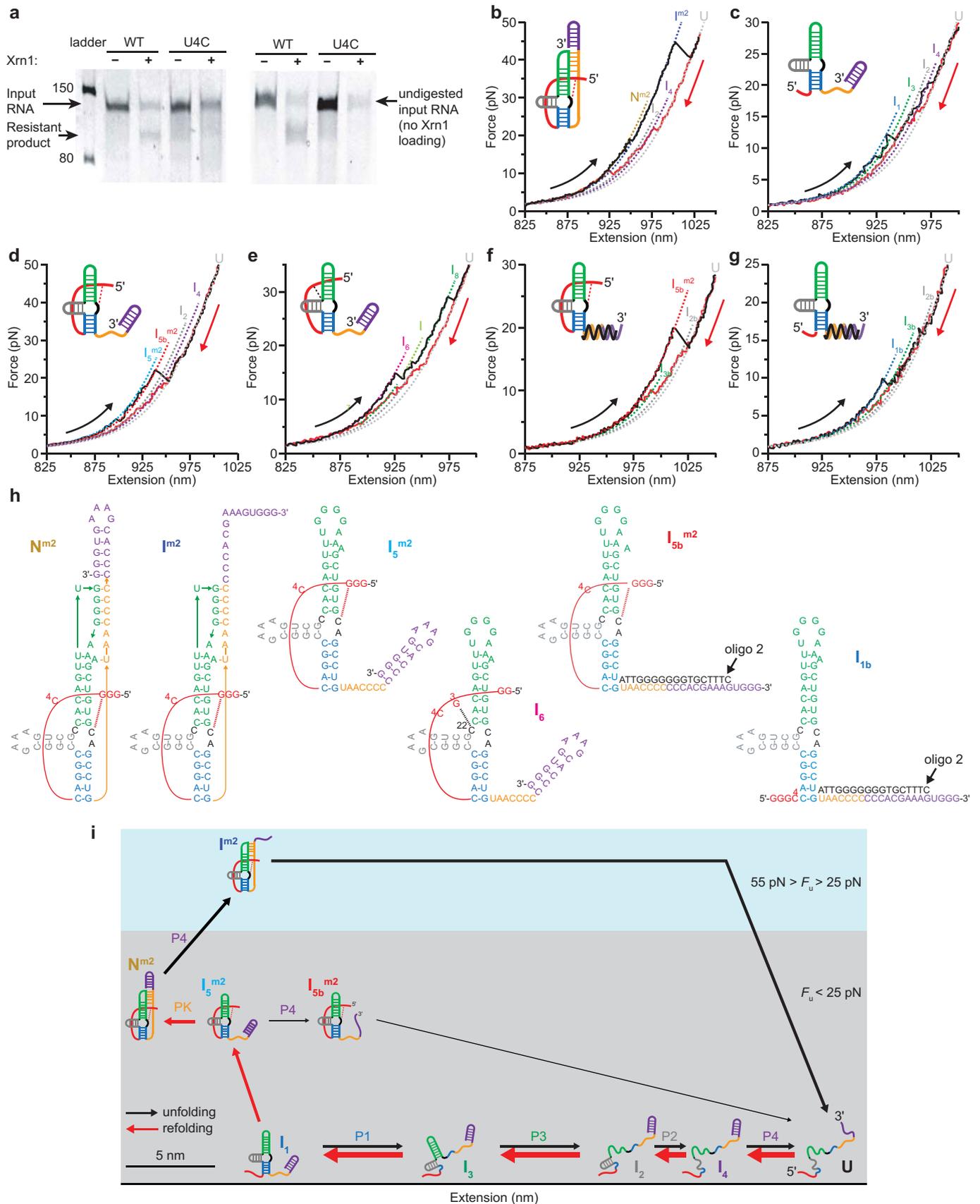
Extended Data Fig. 5 | SMFS of the C22G mutant. **a**, Roughly 15% of unfolding FECs (black) show low forces consistent with secondary structure only, in state I_1' , which is akin to I_1 (Supplementary Fig. 2b) but with shorter P2 and longer P3. Dashed lines show WLC fits to each state, inset shows structure before unfolding. **b**, Roughly 17% of unfolding FECs (black) show a moderately high-force state I_m1 consistent with partial formation of 5'-TC and the non-native pseudoknot PK' but without any 5'-end threading. Refolding FECs (red) show a pathway similar to I_1 (Supplementary Fig. 3a), although with shorter P2 and longer P3, indicating I_m1 derives from I_1' . **c**, Roughly 15% of unfolding FECs (black) show the same extreme stability as for the wild-type ring-knot, indicating this mutant forms a ring-knot analogous to the wild-type knot, with full 5'-TC and extreme mechanical resistance. **d**, Roughly 53% of unfolding (black) and refolding FECs (red) show low forces consistent with state I_1'' (inset), akin to I_1 , containing secondary structure only but with lengthened P2. **e**, Proposed structures of the states identified in (a-d). Base-pairing of G22 with C44 prevents native 5'-TC but allows I_1' - I_4' and I_m1 (with shorter P2 and longer P3) to form. Alternatively, base-pairing of G22 with C9 lengthens P2 but leaves C44 accessible for native 5'-TC formation, allowing formation of mutant states N^{m1} and I_r^{m1} containing a mechanically resistant ring-knot, in addition to intermediates I_1'' , I_2'' , and I_4'' . Contour length changes expected from unfolding each structure are listed in Supplementary Table 4.



Extended Data Fig. 6 | Effects of anti-sense oligo 3 on C22G mutant and evidence for G22:C9 base-pairing. a, b, All (a) unfolding and (b) refolding FECs measured in the presence of oligo 3 show secondary structure only. **c,** Without oligo 3 (red), two sub-populations are seen, with different average contour length changes (top) and unfolding forces (bottom). High-force transitions have lengths inconsistent with expectations for G22:C44 pairing (blue, left) but consistent with expectations for G22:C9 pairing (blue, right). As expected, ΔL_c for unfolding P1+P3 without oligo 3 (Supplementary Table 4) is anti-correlated with the change in ΔL_c for P2 between the two sub-populations, being longer for G22:C44 pairing. With oligo 3 present (black), only one population is seen, matching the results for G22:C9 pairing. F_v is expected to be higher for P2 with G22:C9 pairing because two extra base-pairs are predicted by *mfold* to make P2 ~ 5.7 kcal/mol more stable. Note that these bar graphs restate results presented in Fig. 4c and Supplementary Table 4; error bars represent standard error on the mean. **d,** Proposed structures for intermediate states with oligo 3 present. Contour length changes expected from unfolding each structure are listed in Supplementary Table 4.



Extended Data Fig. 7 | Unfolding and refolding pathways for C22G mutant. The two possible base-pairings for the mutated base (G22), with either C9 or C44, lead to heterogeneous pathways. Superscripts 'e' and 's' denote respectively elongated and shortened versions of a helix relative to the native ones. Arrow thicknesses indicate pathway probabilities.



Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | XR activity and SMFS of U4C mutant. **a**, Measurements of XR activity show U4C mutant has no resistance to digestion by Xrn1 (to within 2% detection limit), in contrast to strong resistance of WT xrRNA (~80%). Note that some of the input RNA remains undigested because of incomplete loading of Xrn1. Two replicates are shown out of 3 independent measurements yielding the same result of ~0% resistance for U4C. **b**, Just over half (52%) of FECs show folding (red) into and unfolding (black) out of a native-like ring-knot. **c**, Most of the remaining FECs (44%) show folding (red) into and unfolding (black) out of intermediate I_1 . **d**, A very few FECs (1%) show folding (red) into and unfolding (black) out of an intermediate with the 5' end natively threaded but the ring still open. **e**, A very few FECs (~3%) for U4C show folding (red) into and unfolding (black) out of a non-native state (I_6). A speculative model of I_6 involving non-native 5'-TC is illustrated (inset, dashed black line: G3:C22), but further study is needed. **f**, With oligo 2 blocking PK, most FECs ($73 \pm 7\%$) unfolded from the threaded intermediate I_{6b}^{m2} , whose unfolding force was lower than for the analogous state in the WT xrRNA (Supplementary Fig. 6a) because of disrupted 5'-TC. **g**, A minority of FECs ($27 \pm 7\%$) with oligo 2 unfolded from I_1 . **h**, Proposed structures of states identified in (b-g). Contour length changes expected from unfolding each structure are listed in Supplementary Table 5. **i**, Pathways for unfolding (black) and refolding (red). Arrow thicknesses indicate pathway probabilities.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- | | | |
|-------------------------------------|-------------------------------------|--|
| <input type="checkbox"/> | <input checked="" type="checkbox"/> | The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | A description of all covariates tested |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
<i>Give P values as exact values whenever suitable.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Data analysis

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	16,447 curves were measured from 129 individual RNA molecules. The number of curves measured from a single molecule before it breaks varies widely, hence sample sizes could not be pre-determined. Approximate minimum sample sizes were chosen to ensure: (i) sufficient sampling to detect intermediates present at the 1% level (i.e., a minimum of 100-200 curves); (ii) sufficient sampling to obtain at least 10 measurements of contour length changes for every intermediate detected (including very rare states), to allow for reliable average values; and (iii) a minimum of 5 molecules for baseline measurements of WT and mutant xrRNAs, or 2 molecules for measurements with anti-sense oligos.
Data exclusions	No data from measurements of single RNA molecules were excluded. All measurements that did not involve correctly assembled single RNA molecules--as reflected via incorrect persistence length, contour length, or overstretching transition force--were excluded.
Replication	Each curve represents an independent measurement of the unfolding/refolding. Measurements on different molecules represent replicates of the experiment. For each experimental condition, 162-5212 independent measurements were done on 2-69 molecules (replicates), as listed in Supplementary Tables 1-5. All the force-extension curves were reproducible across the different molecules (replicates). 3 replicates were done of gel images.
Randomization	Randomization is not relevant to this study since there were no experimental groups. The individual molecules being studied were inherently randomized in solution.
Blinding	There were no experimental groups in this study, hence blinding is not relevant. All curves were analyzed in the same way.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging