




Directional translocation resistance of Zika xrRNA

Antonio Suma ^{1,2}, Lucia Coronel³, Giovanni Bussi ³ & Cristian Micheletti ³✉

xrRNAs from flaviviruses survive in host cells because of their exceptional dichotomic response to the unfolding action of different enzymes. They can be unwound, and hence copied, by replicases, and yet can resist degradation by exonucleases. How the same stretch of xrRNA can encode such diverse responses is an open question. Here, by using atomistic models and translocation simulations, we uncover an elaborate and directional mechanism for how stress propagates when the two xrRNA ends, 5' and 3', are driven through a pore. Pulling the 3' end, as done by replicases, elicits a progressive unfolding; pulling the 5' end, as done by exonucleases, triggers a counterintuitive molecular tightening. Thus, in what appears to be a remarkable instance of intra-molecular tensegrity, the very pulling of the 5' end is what boosts resistance to translocation and consequently to degradation. The uncovered mechanistic principle might be co-opted to design molecular meta-materials.

¹Dipartimento di Fisica, Università di Bari and INFN Sezione di Bari, via Amendola 173, 70126 Bari, Italy. ²Institute for Computational Molecular Science (ICMS), Temple University, 19122 Philadelphia, PA, Italy. ³Physics Area, Scuola Internazionale Superiore di Studi Avanzati (SISSA), Via Bonomea 265, 34136 Trieste, Italy. ✉email: cristian.micheletti@sissa.it

Subgenomic flavivirus RNAs (sfRNAs) are noncoding RNAs that result from the partial degradation of the viral genomic RNAs by cellular Xrn1 exoribonucleases. They accumulate in cells infected by flaviviruses, such as Zika, disrupting several molecular processes of the host. While sfRNAs cannot be degraded by exonucleases, they can still be processed by polymerases and reverse transcriptases^{1–5}.

As part of the ongoing endeavor to explain sfRNAs dichotomic response to processive enzymes, which is at the basis of flaviviruses pathogenicity, a series of targeted *in vitro* assays, including mutagenesis experiments have been carried out. They clarified that resistance to exonucleases is provided by a ca 71-nucleotide-long subsequence, xrRNA in short, that can withstand degradation at its 5′ end by various other exonucleases besides Xrn1⁵. Inspection of xrRNA structure—the one of Zika is shown in Fig. 1—reveals multiple pseudoknots and a ring-like motif that appears well-suited to dock onto the surface of the exonuclease^{2,3}.

The elaborate architecture of the relatively-short xrRNAs poses general questions about the compliance of xrRNAs to being translocated through the lumen of enzymes working from the 5′ to 3′ ends, as done by exonucleases, or from the 3′ to 5′ ones, as done by replicases and reverse transcriptases. What are the mechanistic and thermodynamic underpinnings of xrRNA resistance? Would the same directional resistance be observed in single-molecule setups forcing translocation through narrow pores? And, can one learn general tensegrity principles transferable to molecular design?

Prompted by these questions, we set out to study the microscopic mechanical underpinnings of sfRNAs resistance by using an atomistic model of the Zika xrRNA and stochastic simulations of its driven translocation through a narrow cylindrical pore. The pore-translocation setup provides a convenient abstraction of the resistance of sfRNAs to being driven through various types of exonucleases. In a similar spirit, our model system is exclusively informed with the native three-dimensional organization of the Zika xrRNA, thus discounting sequence-dependent effects and specific interactions with nucleases. More in general, it allows for a comparison with other biologically-motivated contexts where the intra-molecular organization has been investigated in connection with the compliance to translocate through enzymes or synthetic pores^{6–16}.

In our systematic study, we perform hundreds of force-ramped translocation simulations from the 3′ and 5′ ends. Structural effectors of translocation compliance are established by analysing

the network of strained interactions as well as by selectively removing groups of them. Bell–Evans analysis and especially metadynamics simulations are then used to obtain a characterization of the free-energy landscape, of the barriers associated with translocation resistance and of the pathway to the effective transition state.

Throughout the explored combination of ramping protocols and temperatures, we observe that translocation from the 5′ end requires forces that significantly exceed those needed for the 3′ end. This shows that Zika xrRNA architecture, which is the sole input of our model, suffices to encode a significant directional response to translocation. The enhanced hindrance at the 5′ end is shown to originate from an elaborate stress redistribution mechanism that has no analog at the 3′ end. The much higher free-energy barrier encountered at the 5′ end reflects an orders of magnitude difference of the time required to activate translocation from the two ends at constant force.

The results illuminate the structure-based contribution of xrRNA directional resistance and provide a single quantitative framework that recapitulates the resistance to 5′ → 3′ exonuclease degradation and 3′ → 5′ processability by reverse transcriptases and polymerases. It also indicates that pore-translocation setups, which could verify our simulation results, could be used as valuable probes of xrRNA resistance.

Results

Structure of Zika xrRNA. Figure 1a shows the structure of the Zika xrRNA, a 71-nucleotide long noncoding subgenomic RNA that can resist degradation by Xrn1 and other 5′ → 3′ ribonucleases³. Its canonical and noncanonical base pairings are summarized in Fig. 1b. Heavy atoms contacts, including those of sugar and phosphate moieties, are shown in the native contact map of Fig. 1c. The secondary and tertiary organization is articulated over four helices (P1–P4), and multiple pseudoknot-like elements (Pk1–Pk5).

Response to pore translocation. We integrated atomistic modeling, molecular dynamics (MD) simulations, force-spectroscopy analysis, and thermodynamic sampling techniques to understand how exactly the molecular architecture of the Zika xrRNA underpins its resistance to being driven through the lumen of exonucleases from the 5′ end, while still allowing to be processed and unwound by enzymes, such as polymerases and reverse transcriptases, which work from the 3′ end.

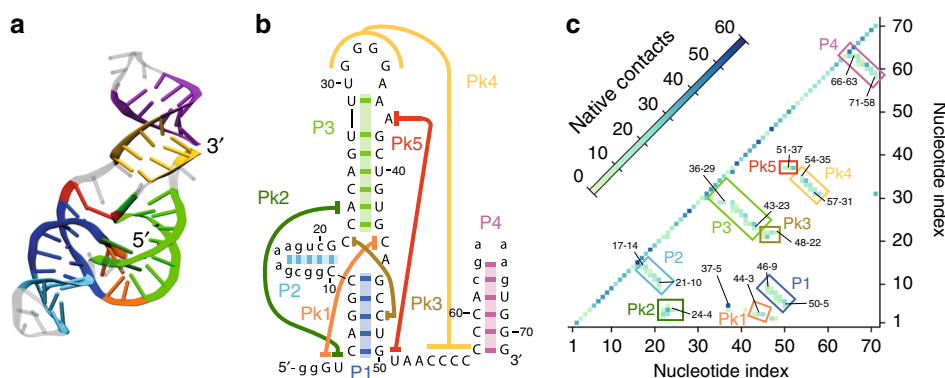


Fig. 1 Structural organization of Zika xrRNA. **a** Ribbon representation of the 71-nucleotide long xrRNA of the Zika virus, PDB entry 5TPY³. **b** Diagrammatic representation of canonical and noncanonical base pairings, as annotated by Barnaba⁴⁷. **c** Heat map of the native contacts. The map includes heavy-atom interactions within a 4 Å cutoff distance and thus include contacts formed by bases, sugar, and phosphate moieties. The four helices, P1–P4, and pseudoknotted regions, Pk1–Pk5, are highlighted. Diagram in **b** follows ref. ³. Secondary elements are highlighted with the same color code across the three panels. Structural renderings are based on the VMD software package⁴⁸.

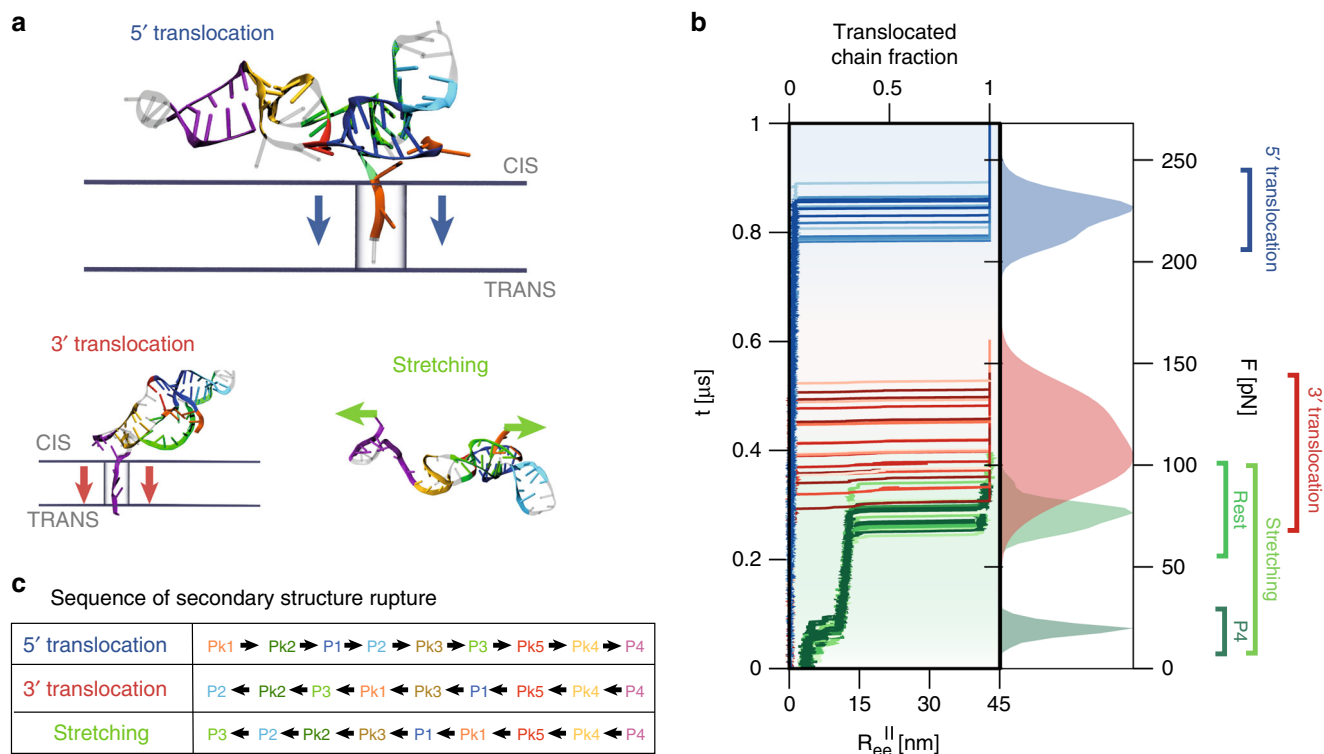


Fig. 2 Zika xrRNA response to pore translocation and stretching. **a** Typical snapshots of the initial stages of translocation and stretching processes; the arrows indicate the direction of the force applied inside the pore or at the chain ends. **b** Temporal trace of the translocated chain fraction from the 5' (blue) and 3' (red) ends for 20 trajectories at the reference temperature $T = 300$ K and force-ramping rate $r = r_0 = 270$ pN/ μ s. The green traces show the longitudinal end-to-end distance, R_{ee}^{\parallel} for 20 stretching trajectories at the same values of T and r . The filled curves on the right are the associated height-normalized distributions of the critical forces for translocation and mechanical unfolding, obtained with Gaussian kernel-density estimators. **c** Typical order of rupture of native contacts during translocation and stretching, see also Supplementary Fig. 2. Secondary elements in **a**, **c** are highlighted with the same color code of Fig. 1a.

Inspired by force-spectroscopy approaches^{17,18}, we adopted a force-ramping protocol where the Zika xrRNA is driven through a cylindrical pore. The pore is 11.7 Å wide and runs perpendicularly through a parallelepiped slab that is 19.5 Å thick, Fig. 2a. The pore geometry approximates the Xrn1 nuclease lumen¹⁹ and allows the passage of only a single RNA strand at a time. Mimicking an electrophoretic setup^{11,20}, the driving force was applied exclusively to the xrRNA *P* atoms inside the pore, whose longitudinal axis we take as the vertical or *z* axis of a Cartesian coordinate system, with the *x*-*y* plane coinciding with the *cis* surface of the slab.

The xrRNA structure was represented with SMOG. The model retains the full atomistic details and has been validated for various molecular systems, including nucleic acids^{21–24}. SMOG employs a native-centric force field that allows us to highlight structure-dependent properties while filtering out details more related to sequence or specific RNA–protein interactions.

We carried out hundreds of ramped translocation simulations of the molecule after priming it at the pore entrance from both ends, Fig. 2a. A representative set of translocation trajectories from 5' and 3' ends is shown in Fig. 2b. The curves show the temporal trace of the translocated fraction of the molecule at the reference temperature $T = 300$ K and force-ramping rate $r_0 = 270$ pN/ μ s. Translocation initiates only after a long loading stage, after which it proceeds precipitously. Major differences in translocation resistance emerge between the 5' and 3' ends.

The unlocking, or triggering of translocation, occurs when the vertical position of the leading *P* atom, measured relative to the *cis* plane, falls below $z = -19.5$ Å or $z = -15$ Å for 5' or 3' entries, respectively, and is followed by a cascade of ruptured

contacts without significant pauses between the breaking of various secondary elements (Supplementary Figs. 2, 3). The typical order of rupture is given Fig. 2c. For the inherent stochasticity of the process, the triggering force is not unique but varies in a range, following a unimodal probability distribution as shown in Fig. 2b. Compared with the 3' case, the most probable triggering force at the 5' end is more than double; it is also much larger than the standard deviation of the triggering forces at the same end, 8 pN, or at the opposite one, 17 pN.

Comparison with xrRNA stretching response. The unusually high resistance at the 5' end is further highlighted by contrasting the translocation forces at this end with those required to unfold the xrRNA by pulling the two termini apart, as in force spectroscopy. As shown in Fig. 2b, unfolding by stretching, which occurs when the longitudinal end-to-end distance, R_{ee}^{\parallel} exceeds 16 nm, requires forces three times smaller than those needed to initiate 5' translocations. Interestingly, the typical sequence of secondary elements rupture by stretching is similar to the case of 3' translocations. The main difference is that the P4 helix, which yields when $R_{ee}^{\parallel} \sim 9$ nm, now unfolds well before the rest of the structure (Fig. 2b and Supplementary Fig. 2).

In summary, at similar conditions, the characteristic force needed to initiate xrRNA translocation from the 5' end, which resists exonuclease degradation, is distinctly higher than the forces that causes unfolding by stretching or by translocation from the 3', which is processable by replicases. The same results are seen for the several hundreds of trajectories gathered over 19 combinations of T and r , from lower to higher temperatures and

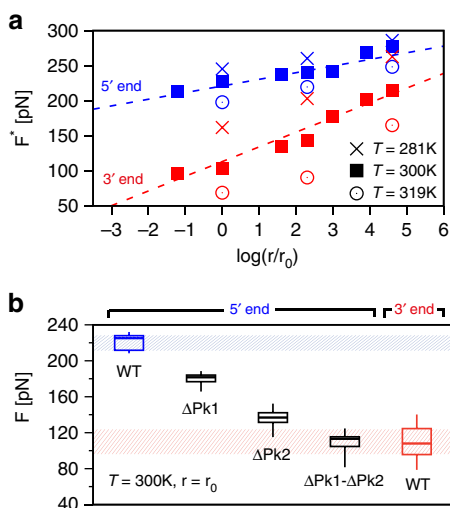


Fig. 3 Critical translocation forces for different conditions and variants of Zika xrRNA. **a** Most probable translocation forces from the 5' and 3' ends at different temperatures, T , and force-ramping rates, r , the estimated errors are smaller than the symbols. For clarity only a subset of all considered (T , r) combinations are shown, see Supplementary Fig. 7. The dashed lines show the Bell-Evans relation of Eq. (1) for $T = 300\text{K}$ with parameters fitted simultaneously on all (T , r) datapoints, see main text. **b** Box-whisker plots (center line, median; box limits, upper and lower quartiles) for the translocation forces at $T = 300\text{K}$ and $r = r_0$, for the native (WT) xrRNA (5' and 3' entries) and for three variants lacking the native contacts in Pk1 and/or Pk2 (5' entries).

faster or slower ramping protocols than the reference values, see Supplementary Figs. 4, 7.

Bell-Evans analysis. The extensive set of translocations at all explored (T , r) combinations can be recapitulated by the Bell-Evans (BE) analysis. In force-spectroscopy contexts it is customary to resort to the following BE expression for the most probable rupture force^{17,18,25}:

$$F^* = \frac{K_B T}{\Delta} \cdot \ln \left(\frac{r \cdot \Delta}{\nu \cdot K_B T} \right) + \frac{E_T}{\Delta}, \quad (1)$$

where ν is a kinetic coefficient and E_T and Δ , the key quantities of interest, are the effective height and width of the thermodynamic barrier associated to the rupture event at zero force.

Figure 3a shows that expression (1) provides a good fit of the most probable translocation forces, yielding a $\chi^2 \sim 3$, hence of order unity, for both the 5' and 3' ends when all 19 (T , r) combinations are considered simultaneously. For clarity, only a subset of the considered datapoints are shown in Fig. 3—see Supplementary Fig. 7 for the complete set—and the best fit to the BE expression is shown for $T = 300\text{K}$ only.

The best-fit parameters are $E_T^{\text{BE}} = 36 \pm 2\text{ kcal mol}^{-1}$ and $\Delta^{\text{BE}} = 4.4 \pm 0.4\text{ \AA}$ for 5' entries and $E_T^{\text{BE}} = 24 \pm 1\text{ kcal mol}^{-1}$ and $\Delta^{\text{BE}} = 2.0 \pm 0.1\text{ \AA}$ for 3' ones. At room temperature, $T = 300\text{K}$, these barriers are approximately equal to $60 K_B T$ and $41 K_B T$ for the 5' and 3' ends, respectively.

These barriers describe the system in the absence of a driving force and reflect that translocation is very unlikely to be activated by thermal fluctuations only, especially at the 5' end. In the presence of a constant driving force, an estimate of the activation time can be obtained via the so-called Bell-Evans lifetimes¹⁸ and the best-fit parameterization of Eq. (1), see “Methods”. For constant driving forces of 50–100 pN, a range accessible to

molecular motors²⁶, the estimated activation time from the 5' end exceeds by more than two orders of magnitude the one at the 3' end.

Directional resistance and stress redistribution. The precipitous unlocking of translocation occurs when the P atom at the 5' end is just about at the *trans* edge of the pore, $z \sim -19.5\text{ \AA}$ (Supplementary Fig. 3). This triggering event is preceded by a long loading phase during which the driving force, which is exclusively imparted to the P atoms inside the pore, propagates an increasing tensile disturbance to the xrRNA that is still on the *cis* side.

The ensuing effects are best discussed by considering how the network of native contacts is distorted during the loading stage. The latter can be quantified in terms of the strain, s , of the native interactions for each nucleotide. Negative or positive values of s indicate that contacting atoms are respectively drawn closer or further apart than in the unstrained xrRNA structure, see “Methods”. Kymoplots of the s profile, during the loading stage of translocation from both ends, are shown in Fig. 4a and Supplementary Fig. 6.

The data and accompanying structures of Fig. 4 clarify that, as the initial transmission of tension along the RNA backbone draws the first nucleotides closer to or into the pore, their network of interacting nucleotides is increasingly distorted. In fact, regions G21-U28 and G38-A45 wrap more tightly around the 5' end, particularly nucleotide G3 (Fig. 4b). Thanks to this tightening, which has no analog at the 3' end, the increasing load is distributed over a broad portion of the molecule, which can hence better withstand the dragging force.

Increasing the load at the 5' end eventually weakens the native interactions in several regions. These include xrRNA portions directly exposed to the traction of the 5' end (U4-G14) and their partner strands (A17-C22 and C47-A52), see Fig. 4c. Note that the distal region A45-C48 is further disrupted by being pressed against the outer surface of the pore (Fig. 4c and Supplementary Fig. 5). At the 3' end, instead, no noticeable tightening occurs around the pulled portion. The terminal tract of the P4 helix and their interacting partners in Pk4 and P4 (G31 and C58-C60) are, instead, directly exposed to the unmitigated pulling into the pore. (Fig. 4a, d). The lack of a mechanism for distributing and dissipating stress is what causes the 3' end, the one processed by replicases, to yield at much lower forces than the 5' end, the one resisting to exonucleases.

The tightened portion at the 5' end, highlighted in blue in Fig. 4, is disconnected along the sequence, but is structurally compact. The emergence of a tightened structural element in response to a pulling force, is a distinctive feature of systems with mechanical tensegrity. In these systems, mechanical resistance results from an equilibrium between elements put under tension (positive strain) and other elements that are put under compression (negative strain) by the same applied force. Thus, the 5' architecture provides a heretofore rare example of intra-molecular tensegrity, where the very pulling of the 5' end is what boosts resistance to translocation and consequently to degradation.

Native contacts network and 5' resistance. We next investigated which set of native contacts and structural elements at the 5' end are most conducive to the redistribution of tensile stress and translocation resistance. Our model allows for directly pinpointing these critical contributors, namely by switching off selected combinations of native secondary contacts from the interaction potential. In experiments this is not possible, and similar information can be obtained by performing mutation scans where the identity of nucleotides is affected.

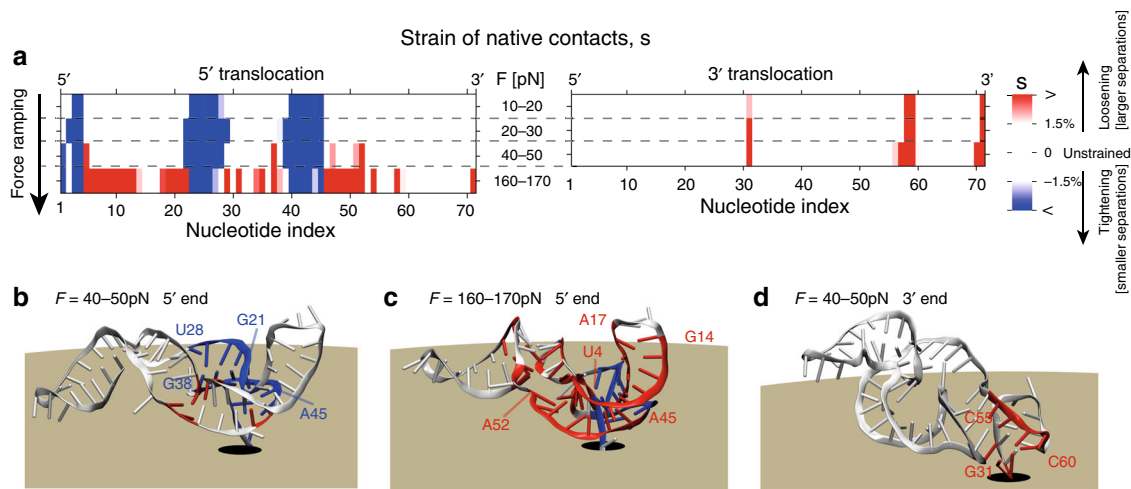


Fig. 4 Mechanics of Zika xrRNA directional resistance to translocation. **a** Kymograph of the strain of native interactions, s , during the loading stages of force-ramped translocation. Negative or positive value of the strain parameter indicate that natively-interacting atoms are, respectively, closer or further apart than in unstrained xrRNA, reflecting the local tightening or loosening of the structure. The nucleotide-wise profile of the strain was computed for the indicated force ranges over 20 translocation trajectories at $T = 300$ K and $r = r_0$, see “Methods” and Supplementary Fig. 6. Nucleotides whose native contact network is significant strained ($|s| > -1.5\%$) are highlighted in **b-d**. The color code is the same as for **a**.

Our analysis indicates that the observed resistance is virtually ascribable to the sole contacts in pseudoknots Pk1 and Pk2. These, we recall, include contacts formed by sugar and phosphate moieties, Fig. 1c. The results are illustrated in Fig. 3b as box-whisker plots for the critical translocation forces of the native xrRNA chain, and three variants lacking the native contacts in either of the two pseudoknots, Δ Pk1 and Δ Pk2, or both, Δ Pk1-Pk2.

In variant Δ Pk1 the critical force is appreciably lowered, but it still is almost twice as large as for the 3' end. A stronger variation is seen for Δ Pk2, whose critical force is reduced so much that it overlaps with the critical forces for the 3' end. A complete loss of directional resistance is found for Δ Pk1-Pk2, which yields practically the same forces encountered at the 3' end. This implies a fallback to a baseline hindrance once Pk1 and Pk2 are disrupted. Of the two sets of pseudoknotted contacts, those in Pk2 are noticeably the most important ones.

5' resistance, free-energy profile and transition state. We finally used the metadynamics framework²⁷ to characterize 5' resistance in more detail and beyond the BE analysis of out-of-equilibrium trajectories. In metadynamics, reversible transitions across significant barriers are achieved with an affordable computational expenditure by using history-dependent biases on multiple order parameters at the same time. Building on the centrality of Pk1 and Pk2 contacts on translocation resistance (Fig. 3b) we used the corresponding fraction of native contacts, Q_{Pk1} and Q_{Pk2} , as natural order parameters, together with the amount of pore insertion of the 5' P atom, z . To avoid effects related to the refolding of the RNA portion that has passed to the *trans* side, we extended the length of the pore and profiled the free energy up to $z = -22$ Å.

A first result of the metadynamics analysis is the dominant translocation pathway shown in Fig. 5a. The pathway projections in the (Q_{Pk1}, z) and (Q_{Pk2}, z) planes clarify that during the initial stages of translocation, z from 0 to -10 Å, contacts in Pk1 and Pk2 are only modestly disrupted. Further insertion into the pore, z from -10 Å to -15 Å, causes an appreciable loss of Pk2 contacts. The final insertion tract, z from -15 Å to -19.5 Å, produces a concurrent loss of contacts in both Pk1 and Pk2. Representative structures along the pathway to the force-induced transition state are shown next to the pathway, along with two-

dimensional projections of the free-energy profile at the same z values, Fig. 5b-d.

The results clarify that Pk2 interactions yield at $z \sim -14$ Å, when the contacts involving the sugar and phosphate moieties of nucleotides G3-C22 are sheared and disrupted by the drawing of G3 into the pore while C22 is sustained and held in place by the elaborate secondary and tertiary architecture surrounding the 5' end. The disruption of the nonlocal contact G3-C22 opens the way for the disruption of other contacts in Pk2 and also Pk1, such as G3 with C44 and A45, for $z \sim -18$ Å.

We note that the dominant pathway established from metadynamics analysis has direct bearings on out-of-equilibrium, force-ramped translocations too. This is clarified by the scattered datapoints in Fig. 5a which represent the same order parameters for conformations sampled from the force-ramped translocation trajectories at $T = 300$ K and $r = r_0$. The agreement between the dominant pathway and the trajectories has three implications: it corroborates both methods, it indicates that translocation-induced unfolding occurs via a single pathway, and shows that the latter is the same in and out-of-equilibrium. The results thus cover slower ramping rates than considered in Fig. 3, including those accessible experimentally.

The salient thermodynamics of the process is aptly conveyed by the one-dimensional free-energy profile of Fig. 5e, obtained by thermodynamic integration of the metadynamics free-energy at fixed z . The free-energy increase up to $z \sim -10$ Å reflects the loss of conformational entropy and native contacts during the loading stage. The steeper rise beyond $z \sim -10$ Å reflects the progressive disruption of Pk2 and Pk1 contacts. The free-energy cost for reaching the depth $z = -19.5$ Å, the triggering condition in force-ramped simulations, is about 46 kcal mol⁻¹.

Applying constant translocation forces tilts the free-energy profile as shown in Fig. 5e. One notes that at sufficiently high force the slope of the free-energy changes sign at about $z \sim -20$ Å. This value of z pleasingly matches the aforementioned triggering insertion depth obtained in force-ramp translocations (Supplementary Fig. 3). The pore insertion depth $z \sim -20$ Å can be interpreted as an effective transition state, embodying the rate-limiting step for force-induced translocations.

In summary, the metadynamics results aptly complement those from force-ramped translocations, giving detailed indications of

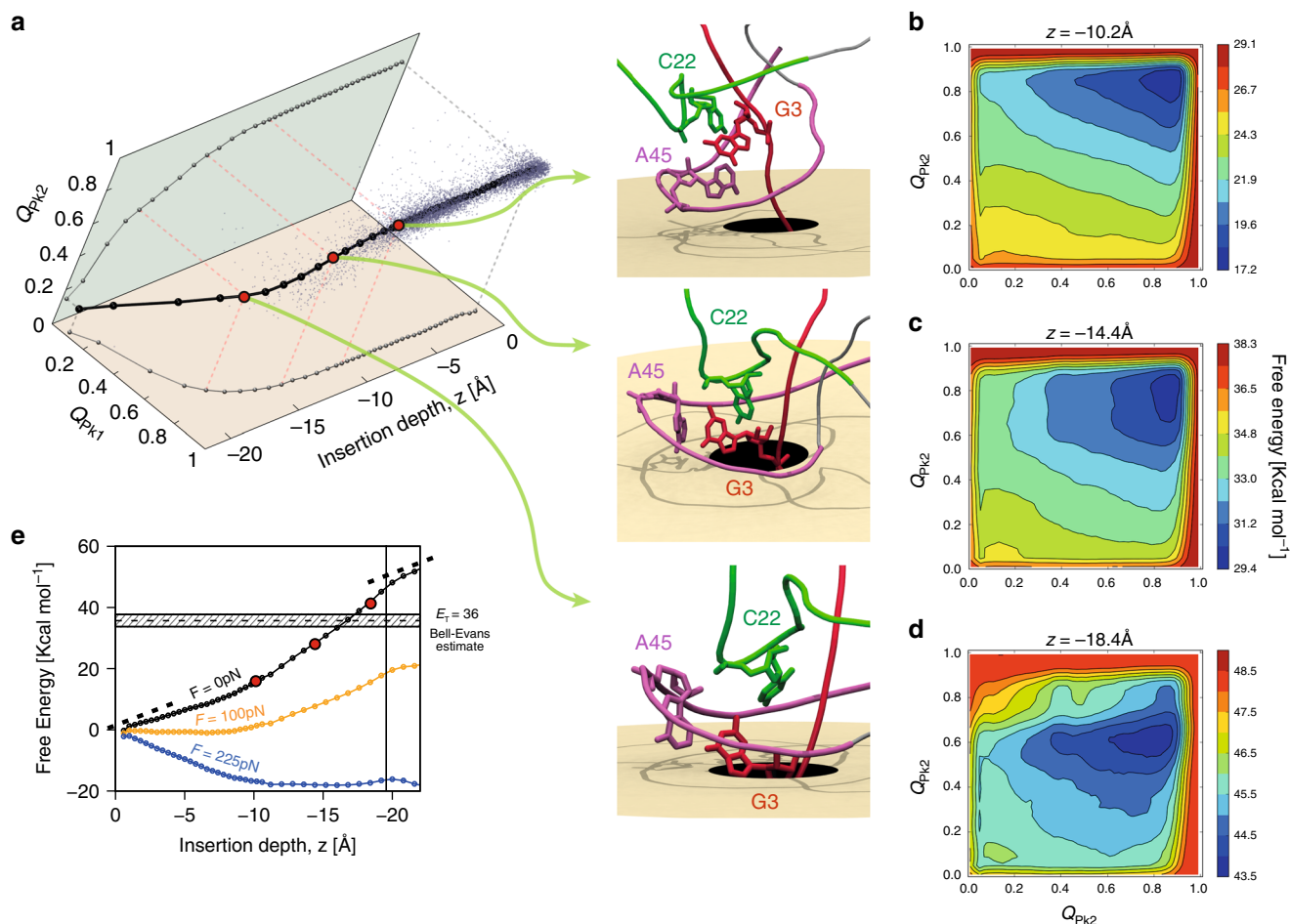


Fig. 5 Free-energy profiling of Zika xrRNA translocation. **a** Dominant translocation pathway, from metadynamics analysis, showing the fraction of native contacts, Q_{Pk1} and Q_{Pk2} , versus z , the pore insertion depth of the 5' P atom. Its projections in the (Q_{Pk1}, z) and (Q_{Pk2}, z) planes are also shown. The cloud of points shows instantaneous values of the same order parameters in force-ramped translocations at $T = 300$ K and $r = r_0$. Representative snapshots at marked values of z (red dots) are shown. **b–d** Sections of the three-dimensional metadynamics free-energy at the same marked values of z . **e** One-dimensional free-energy profile, obtained from thermodynamic integration and reweighting at the indicated translocation forces. The estimated uncertainty is about $0.55 \text{ kcal mol}^{-1}$.

the free-energy barrier associated with the triggering of translocation from the 5' end, and of the associated pathway. The combined analysis shows that, in the loading stage, regions G21–U28 and G38–A45, which surround the 5' end, become more tightly wrapped around it. This mechanism, that has no analog at the 3' end, stabilizes the network of long-range native contacts around nucleotide G3, hindering their entrance into pore until contacts in Pk1 and Pk2, including G3–C22, are eventually disrupted.

Discussion

Nanopore translocation is a powerful method to probe biopolymers as well as understanding how they interact with processive enzymes^{8,15,24,28–33}. Here we used it to clarify the structural determinants of xrRNAs resistance to exonuclease degradation. To this end we carried out hundreds of simulations where an atomistic, native-centric model of Zika xrRNA was translocated through narrow a pore from either end. The setup provides a transparent abstraction of the biological process where the 5' and 3' xrRNA termini are engaged and driven through the lumen of enzymes, such as nucleases or polymerases, encountering different directional resistance.

First, we observed that translocation resistance is significantly higher at the 5' end, the one resisting exonuclease degradation, than at the 3' one, the one yielding to the unwinding operated by

replicases and reverse transcriptases. This holds systematically throughout the many considered combinations of temperatures and force-ramping protocols. It should not go unnoticed that, in our ramping simulations, 3'-entry translocations not only initiate, but also complete at forces well below those needed to start translocation from the 5' end. The xrRNA resistance to unfold when pulled through a pore by the 5' end was also shown to be significantly higher than to unfold by pulling away the two ends (stretching). Bell–Evans analysis indicates that, in the absence of a driving force, the translocation free-energy barriers at the two ends differ by about 12 kcal mol^{-1} . At constant forces of 50–100 pN, the estimated difference of activation times exceeds two orders of magnitude, and ought to be detectable experimentally. Incidentally, we note that the directional effect, as well as the different response to translocation and stretching, suggest that experiments where the two ends are pulled apart might lead to results that do not report directly on the resistance offered to electrophoretic translocation and processive molecular motors.

Secondly, the directional effect originates from the different ways in which the pulling effects are propagated from the termini to the rest of the molecule. During the early loading stages, pulling the 5' end causes regions G21–U28 and G38–A45 to wrap more tightly around it. At later loading stages, interactions

between the pulled terminal tract U4-G14 and partner strands A17-C22 and A45-A52 start to yield. At the same time, region A45-C48 also distorts due to the tight contact with the pore surface. Both mechanisms reflect in a compression of the region surrounding the pulled 5' terminal; this region is compact in space but not in sequence. The tension-induced compression, and the ensuing resistance, make the effect entirely analogous to that of systems with mechanical tensegrity, which oppose tensile forces with compressive ones elicited by the applied tension itself.

As in tensegrity systems, the resistance of the 5' end is encoded in the architectural organization of the molecule and how it interacts with the pore. By contrast, no such mechanism is available at the 3' end for its simpler architecture. By selectively removing native secondary interactions, we established that 5' translocation resistance is primarily encoded in pseudoknot Pk2, and secondarily in Pk1. Once Pk1 and Pk2 contacts are disrupted, no other outstanding barriers are encountered at the 5' end, and the resistance becomes comparable to that of the 3' end. The latter fact also rules out the helical asymmetry of nucleic acids, see e.g., refs. 23,34, as a cause for xrRNAs directional resistance.

Lastly, analysis of free-energy calculations and out-of-equilibrium simulations are consistent at indicating that translocation and the concomitant unfolding from the 5' end occur via a single dominant pathway. Inspection of the pathway shows that a critical step is the disruption of the interactions of nucleotides G3 and C22, which unlocks the breaking of the other Pk2 and Pk1 contacts eventually triggering the irreversible translocation. Also, comparison with metadynamics results indicates that free-energy barriers estimated with customary Bell–Evans analysis are susceptible to simplifying assumptions on the barrier widths. This finding might be relevant in the analysis of force-ramp experiments too.

In conclusion, these results demonstrate that the structural architecture of xrRNA, which is the sole feature informing our model, is predisposed to generate a profoundly asymmetric resistance to translocation from the two ends. This occurs because mechanical stress is distributed and accumulated in the network of contacts in a way that depends not only on the end that is directly pulled, but also on the regions interacting with it, that become more tightly wound around it. This pulling-induced tightening protects the 5' end and allows it to better withstand further pulling of the end itself.

Note that this resistance mechanism is altogether different from the catch-bonding behavior³⁵ of knotted molecules translocating through narrow pores, which is qualitatively opposite to what observed here. Knotted filaments, particularly those with twist knots, can translocate at sufficiently low forces but jam at high ones³⁶. For such systems, jamming happens when the knot, which is pulled against the pore entrance, becomes tight enough that any stochastic sliding motion of the chain along its knotted contour is suppressed^{9,36}. In xrRNA, instead, translocation is stalled at low forces but is unjammed at sufficiently high ones, and moreover in a directional-dependent manner. For xrRNA the initial resistance is offered by the strained network of native interactions that compactify the molecule around the region at the pore entrance, but that eventually yields at sufficiently high forces.

The results help rationalize and recapitulate in a single structural and quantitative framework not only xrRNA resistance to 5' degradation from various exonucleases⁵, but also the concomitant much lower hindrance to being fully processable from the 3' in order to be copied by virally encoded RNA-dependent RNA polymerases. Both properties thus find a natural explanation as different manifestations of the same architecture-dependent response. The results also indicate that present-day experimental setups for pore translocation ought to be adequate to

capture and reveal the discussed signature effects of the xrRNA directional resistance. In perspective, the structural organization principles that Zika xrRNAs have acquired by evolutionary selection could be co-opted for designing synthetic molecules or meta-materials³⁷ with directional resistance to translocation and context-dependent tensegrity response.

Methods

Zika xrRNA and system setup. The reference, native conformation of Zika xrRNA was taken from the protein databank³⁸ entry 5TPY³. Simulations of the 71-nucleotide long RNA (1527 heavy atoms) were based on SMOG, a native-centric model retaining the full atomistic structural details^{39,40}. The SMOG potential includes terms for bonded, nonbonded, angular, and dihedral interactions. The model is designed for exposing structurally-encoded properties by neglecting effects, such as transient formation of non-native contacts and details of interaction with water and ions. These effects are not expected to be crucial in this specific mechanistic application where, besides, the faster ramping protocols called by explicit solvent treatments would antagonize directional resistance, Fig. 3a. Default values for the relative strength of the potentials and the cutoff contact map were used^{40,41}. The absolute temperature was calibrated similarly to ref. 23 by matching the experimental condition that mechanical unfolding of RNA helices^{42,43} occur at about 15 pN at 300 K, see Supplementary Methods and Supplementary Table 1.

The P atom at the 5' or 3' end was primed at the entrance of a cylindrical pore and prevented from retracting from it by a ratchet-like restraint. The pore is 11.7 Å wide and transverses perpendicularly a 19.5 Å thick slab, which separates the *cis* and *trans* sides. The pore geometry approximates the lumen of the Xrn1 exoribonucleases, PDB entry 2Y35¹⁹. The same SMOG inter-atom excluded volume potential was used for steric interactions of the RNA atoms with the surface of the slab and pore.

Molecular dynamics simulations. Langevin MD simulations were carried out with the LAMMPS package⁴⁴, with proper atomic masses, default values for the friction coefficient and with integration timestep equal to $8.7 \times 10^{-4} \tau_{MD}$, where τ_{MD} is the characteristic MD timescale, see Supplementary Methods. Chain translocation from the *cis* to the *trans* sides are driven by a force, F , parallel to the pore axis, and that is exclusively applied to the P atoms inside the pore, among which it is equally subdivided at any given time. The driving force is ramped linearly with time t , $F = r \cdot t$. In stretching simulations, the 5' and 3' P atoms are pulled in opposite directions with the same loading protocol. We considered seven equispaced temperature in the range 281–319 K, and as many pulling protocols, $r/r_0 = 0.3, 1, 5, 10, 20, 50, 100$, with $r_0 = 6.2 \times 10^{-4}$ pN/ τ_{MD} . Conventional mapping to physical units yields $\tau_{MD} \sim 2.3$ ps, see Supplementary Methods, which yields $r_0 = 270$ pN/ μ s, which is conservative for molecular dynamics simulations though still faster than experiments. Note, however, due to the lack of explicit solvent in the model, the effective simulated timescale (pulling rate) reported here represents a lower (upper) bound estimate⁴⁵, implying a better agreement with experimental pulling rates. A total of 19 combinations of parameters were explored, see Supplementary Table 2, and for any of them 20 translocation simulations were typically collected for each orientation (5' → 3' and vice versa). At the reference temperature and ramping rate ($T = 300$ K, $r = r_0$) the number of timesteps needed to trigger translocation from the 5' end is about 7×10^8 . Including stretching simulations, more than 1000 trajectories were run on the Ulysses supercomputer at SISSA.

Bell–Evans analysis. For each orientation, the most probable translocation forces at the 19 combinations of temperature and loading rate, $F^*(T, r)$, were simultaneously fitted with the Bell–Evans expression of Eq. (1). The fit consisted of a χ^2 minimization over the free parameters E_T , Δ and v . $F^*(T, r)$ was established from a Gaussian kernel-density analysis of translocation triggering events, corresponding to a pore insertion depth of the terminal P atom equal to $z = -19.5$ Å and -15 Å for 5' and 3' entries, respectively. The estimated error on $F^*(T, r)$ was set equal to the error of the mean translocation force. The BE waiting time for triggering translocation at a fixed force, F , is $\tau = F_\beta/r_F$, where $F_\beta = K_B T/\Delta$ and r_F is the loading rate for which $F^* = F$ ¹⁸.

Strain analysis. We considered the relative strain of a nucleotide, $s = (q^0 - q^F)/q^0$. q^0 is the average contact fraction (overlap) of all native interactions of the nucleotide computed at a reference range of low forces, 0–10 pN. q^F is analogously computed at force F . The contact fraction of two atoms at distance d is $q(d) = 1/(1 + [d/(1.5 \cdot d^0)]^6)$, d^0 being their native distance (default cutoff distance of 4 Å).

Metadynamics. Metadynamics simulations²⁷ were carried out with the COLVAR module⁴⁶ using three collective variables: z , Q_{Pk1} and Q_{Pk2} , which are respectively the depth of insertion of the 5' P atom into the pore, and the fraction of native contacts in pseudoknots Pk1 and Pk2. Q_{Pk} was computed as $\frac{1}{n_c} \sum_{j=1}^{n_c} q(d_j)$, where j runs over the n_c native contacts of the pseudoknot. Unbiased simulations, see Supplementary Fig. 8,

were used to set the widths of the metadynamics multivariate Gaussians ($\sigma_z = 0.25 \text{ \AA}$, $\sigma_{Q_{PK1}} = \sigma_{Q_{PK2}} = 0.0125$), their height ($0.15 \text{ kcal mol}^{-1}$) and deposition interval (10^3 timesteps). Retraction of chain from the pore was prevented by a restraining potential, $V = \frac{k}{2} \cdot z^2 \cdot \theta(-z)$, where $k = 15.6 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ and θ is the Heaviside function. For the metadynamics analysis, the pore was extended so to reconstructs the free-energy landscape up to $z = -22 \text{ \AA}$, which required runs of 1.5×10^9 timesteps, Supplementary Fig. 9. A constant force bias of 110 pN was used to reduce the time required for filling the basin. The bias was then discounted with a suitable reweighting *a posteriori*. The dominant pathway was obtained by discretizing z with intervals of width $\Delta z = 0.8 \text{ \AA}$ and computing the corresponding canonical averages of Q_{PK1} and Q_{PK2} .

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

Data availability

Data supporting the findings of this paper are available from the corresponding author upon reasonable request. A reporting summary for this Article is available as a Supplementary Information file. The source data underlying Figs. 1c, 2b, 3a, b, 4a, 5 are provided as Source data file. Source data are provided with this paper.

Code availability

Typical setup files for the LAMMPS simulations are provided as Supplementary Software 1. Setup files for SMOG simulations, natively created for GROMACS, were converted for LAMMPS using custom extensions and conversion scripts. The latter are made freely available for academic use at the "SMOG-converter" GitHub repository, github.com/CristianMicheletti/SMOG-converter. Source data are provided with this paper.

Received: 16 March 2020; Accepted: 29 June 2020;

Published online: 27 July 2020

References

- 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).
- Chapman, E. G. et al. The structural basis of pathogenic subgenomic flavivirus RNA (sfRNA) 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).
- Slonchak, A. & Khromykh, A. A. Subgenomic flaviviral RNAs: what do we know after the first decade of research. *Antivir. Res.* **159**, 13–25 (2018).
- MacFadden, A. et al. Mechanism and structural diversity of exoribonuclease-resistant RNA structures in flaviviral RNAs. *Nat. Commun.* **9**, 119 (2018).
- Luo, K., Ala-Nissila, T., Ying, S.-C. & Bhattacharya, A. Sequence dependence of DNA translocation through a nanopore. *Phys. Rev. Lett.* **100**, 058101 (2008).
- 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. U.S.A.* **106**, 12706–12711 (2009).
- Comer, J., Dimitrov, V., Zhao, Q., Timp, G. & Aksimentiev, A. Microscopic mechanics of hairpin DNA translocation through synthetic nanopores. *Biophys. J.* **96**, 593–608 (2009).
- Suma, A., Rosa, A. & Micheletti, C. Pore translocation of knotted polymer chains: how friction depends on knot complexity. *ACS Macro Lett.* **4**, 1420–1424 (2015).
- Ziegler, F. et al. Knotting and unknotting of a protein in single molecule experiments. *Proc. Natl Acad. Sci. U.S.A.* **113**, 7533–7538 (2016).
- Plesa, C. et al. Direct observation of DNA knots using a solid-state nanopore. *Nat. Nanotechnol.* **11**, 1093–1097 (2016).
- Suma, A. & Micheletti, C. Pore translocation of knotted DNA rings. *Proc. Natl Acad. Sci. U.S.A.* **114**, E2991–E2997 (2017).
- San Martín, Á. et al. Knots can impair protein degradation by atp-dependent proteases. *Proc. Natl Acad. Sci. U.S.A.* **114**, 9864–9869 (2017).
- Sriramoju, M. K., Chen, Y., Lee, Y.-T. C. & Hsu, S.-T. D. Topologically knotted deubiquitinases exhibit unprecedented mechanostability to withstand the proteolysis by an aaa+ protease. *Sci. Rep.* **8**, 7076 (2018).
- Cao, C. et al. Mapping the sensing spots of aerolysin for single oligonucleotides analysis. *Nat. Commun.* **9**, 1–9 (2018).
- Sivertsson, E. M., Jackson, S. E. & Itzhaki, L. S. The aaa+ protease clpXP can easily degrade a 3 1 and a 5 2-knotted protein. *Sci. Rep.* **9**, 1–14 (2019).
- Dudko, O. K., Hummer, G. & Szabo, A. Theory, analysis, and interpretation of single-molecule force spectroscopy experiments. *Proc. Natl Acad. Sci. U.S.A.* **105**, 15755–15760 (2008).
- Evans, E. Probing the relation between force-lifetime-and chemistry in single molecular bonds. *Annu. Rev. Biophys.* **30**, 105–128 (2001).
- 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).
- Kowalczyk, S. W., Wells, D. B., Aksimentiev, A. & Dekker, C. Slowing down DNA translocation through a nanopore in lithium chloride. *Nano Lett.* **12**, 1038–1044 (2012).
- Whitford, P. C. & Sanbonmatsu, K. Y. Simulating movement of tRNA through the ribosome during hybrid-state formation. *J. Chem. Phys.* **139**, 121919 (2013).
- Nguyen, K. & Whitford, P. C. Steric interactions lead to collective tilting motion in the ribosome during mRNA-tRNA translocation. *Nat. Commun.* **7**, 1–9 (2016).
- Colizzi, F. et al. Asymmetric base-pair opening drives helicase unwinding dynamics. *Proc. Natl Acad. Sci. U.S.A.* **116**, 22471–22477 (2019).
- Bandarkar, P., Yang, H., Henley, R. Y., Wanunu, M. & Whitford, P. C. How nanopore translocation experiments can measure RNA unfolding. *Biophys. J.* **118**, 1612–1220 (2020).
- Björnham, O. & Andersson, M. Theory for nonlinear dynamic force spectroscopy. *Eur. Biophys. J.* **46**, 225–233 (2017).
- Smith, D. E. et al. The bacteriophage $\phi 29$ portal motor can package DNA against a large internal force. *Nature* **413**, 748–752 (2001).
- Laio, A. & Parrinello, M. Escaping free-energy minima. *Proc. Natl Acad. Sci. U.S.A.* **99**, 12562–12566 (2002).
- Kantor, Y. & Kardar, M. Anomalous dynamics of forced translocation. *Phys. Rev. E* **69**, 021806 (2004).
- Branton, D. et al. The potential and challenges of nanopore sequencing. *Nat. Biotechnol.* **26**, 1146–1153 (2008).
- Henley, R. Y., Carson, S. & Wanunu, M. Studies of RNA sequence and structure using nanopores. *Prog. Mol. Biol. Transl. Sci.* **139**, 73–99 (2016).
- Meller, A., Nivon, L. & Branton, D. Voltage-driven DNA translocations through a nanopore. *Phys. Rev. Lett.* **86**, 3435 (2001).
- Trepagnier, E. H., Radenovic, A., Sivak, D., Geissler, P. & Liphardt, J. Controlling DNA capture and propagation through artificial nanopores. *Nano Lett.* **7**, 2824–2830 (2007).
- Cao, C., Yu, J., Wang, Y.-Q., Ying, Y.-L. & Long, Y.-T. Driven translocation of polynucleotides through an aerolysin nanopore. *Anal. Chem.* **88**, 5046–5049 (2016).
- Mathé, J., Aksimentiev, A., Nelson, D. R., Schulten, K. & Meller, A. Orientation discrimination of single-stranded DNA inside the α -hemolysin membrane channel. *Proc. Natl Acad. Sci. U.S.A.* **102**, 12377–12382 (2005).
- Pereverzev, Y. V., Prezhdo, O. V., Forero, M., Sokurenko, E. V. & Thomas, W. E. The two-pathway model for the catch-slip transition in biological adhesion. *Biophys. J.* **89**, 1446–1454 (2005).
- Rosa, A., Di Ventra, M. & Micheletti, C. Topological jamming of spontaneously knotted polyelectrolyte chains driven through a nanopore. *Phys. Rev. Lett.* **109**, 118301 (2012).
- Paulose, J., Chen, B. G.-g. & Vitelli, V. Topological modes bound to dislocations in mechanical metamaterials. *Nat. Phys.* **11**, 153–156 (2015).
- Berman, H. M., Bourne, P. E., Westbrook, J. & Zardecki, C. The protein data bank. In *Protein Structure*, 394–410 (CRC Press, 2003).
- Whitford, P. C. et al. An all-atom structure-based potential for proteins: bridging minimal models with all-atom empirical forcefields. *Proteins* **75**, 430–441 (2009).
- Noel, J. K. et al. Smog 2: a versatile software package for generating structure-based models. *PLOS Comput. Biol.* **12**, e1004794 (2016).
- Hayes, R. L. et al. Reduced model captures Mg²⁺-RNA interaction free energy of riboswitches. *Biophys. J.* **106**, 1508–1519 (2014).
- Manosas, M., Collin, D. & Ritort, F. Force-dependent fragility in RNA hairpins. *Phys. Rev. Lett.* **96**, 218301 (2006).
- Wen, J.-D. et al. Force unfolding kinetics of RNA using optical tweezers. *Biophys. J.* **92**, 2996–3009 (2007).
- Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **117**, 1–19 (1995).
- Yang, H. et al. Diffusion of tRNA inside the ribosome is position-dependent. *J. Chem. Phys.* **151**, 085102 (2019).
- Fiorin, G., Klein, M. L. & Hémin, J. Using collective variables to drive molecular dynamics simulations. *Mol. Phys.* **111**, 3345–3362 (2013).
- Bottaro, S. et al. Barnaba: software for analysis of nucleic acid structures and trajectories. *RNA* **25**, 219–231 (2019).
- Humphrey, W., Dalke, A. & Schulten, K. Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).

Acknowledgements

This work was partially supported by MIUR, the Italian Ministry of Education. We thank Antonio De Simone for useful discussions.

Author contributions

A.S., L.C., G.B., and C.M. designed research; A.S., L.C., and C.M. carried out the simulations; A.S. and C.M. did the analysis and wrote the paper with inputs from G.B.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41467-020-17508-7>.

Correspondence and requests for materials should be addressed to C.M.

Peer review information *Nature Communications* thanks Aleksei Aksimentiev, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information is available at <http://www.nature.com/reprints>

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020