Article

Stick-slip unfolding favors self-association of expanded *HTT* mRNA

Brett M. O'Brien[1], Roumita Moulick[2], Gabriel Jiménez-Avalos ®[2], Nandakumar Rajasekaran[3], Christian M. Kaiser[3,4,5] ✉ & Sarah A. Woodson ®[1,2,5] ✉

In Huntington's Disease (HD) and related disorders, expansion of CAG trinucleotide repeats produces a toxic gain of function in affected neurons. Expanded *huntingtin* (exp*HTT*) mRNA forms aggregates that sequester essential RNA binding proteins, dysregulating mRNA processing and translation. The physical basis of RNA aggregation has been difficult to disentangle owing to the heterogeneous structure of the CAG repeats. Here, we probe the folding and unfolding pathways of exp*HTT* mRNA using single-molecule force spectroscopy. Whereas normal *HTT* mRNAs unfold reversibly and cooperatively, exp*HTT* mRNAs with 20 or 40 CAG repeats slip and unravel non-cooperatively at low tension. Slippage of CAG base pairs is punctuated by concerted rearrangement of adjacent CCG trinucleotides, trapping partially folded structures that readily base pair with another RNA strand. We suggest that the conformational entropy of the CAG repeats, combined with stable CCG base pairs, creates a stick-slip behavior that explains the aggregation propensity of exp*HTT* mRNA.

Huntington's Disease (HD), a fatal neurodegeneration, is caused by expansion of the CAG repeats located in the first exon of the *huntingtin* (*HTT*) gene, in which disease severity and age of onset strongly correlate with the number of CAG repeats[1]. Aggregation of the polyglutamine-containing HTT protein has been extensively studied[2]. Emerging evidence, however, suggests that toxicity of the expanded (exp) *HTT* mRNA also contributes to disease[3–5]. Firstly, transcription is neurotoxic even when the RNA cannot be translated[6]. Secondly, a genome wide association study revealed that the length of the uninterrupted CAG repeat tract, rather than polyQ tract length, is the primary determinant of HD age of onset[7,8]. Thirdly, cellular models of HD have shown that exp*HTT* mRNA self-associates into nuclear or perinuclear aggregates that sequester the splicing regulatory protein, MBNL1, and other RNA-binding proteins[9–13]. These aggregates are thought to arise from the inherent propensity of RNAs containing expanded CAG tracts to phase separate into condensates[11,14,15].

Physical properties that promote RNA aggregation and protein sequestration have been difficult to study because the repetitive sequences adopt many structures. RNAs with 12-15 CAG trinucleotide repeats form hairpins containing A-A mismatches[16,17] that may "slip" into alternative alignments of the opposing strands, exposing a variable number of single-stranded triplets. In simulations, these unpaired CAG triplets seed inter-strand base pairs[18], explaining why phase separation propensity increases with the number of CAG triplets[11,19].

Sequences flanking the CAG repeats in *HTT* also affect mRNA structure and may contribute to repeat expansion toxicity. Chemical probing experiments showed that 7-9 flanking CCG proline codons in *HTT* mRNA base pair with CAG triplets, forming a composite hairpin at the base of the stem[10,20] (Fig. 1A). The CCG repeats increase the melting temperature of the RNA by ~15 °C compared to pure CAG repeats of identical length[10]. It is unknown how CAG-CCG base pairs alter interactions between RNA strands.

[1]Chemical Biology Interface Program, Johns Hopkins University, Baltimore, MD, USA. [2]T. C. Jenkins Department of Biophysics, Johns Hopkins University, Baltimore, MD, USA. [3]Department of Biology, Johns Hopkins University, Baltimore, MD, USA. [4]Bijvoet Center for Biomolecular Research, Utrecht University, Utrecht, The Netherlands. [5]These authors jointly supervised this work: Christian M. Kaiser, Sarah A. Woodson. ✉e-mail: c.m.kaiser@uu.nl; swoodson@jhu.edu
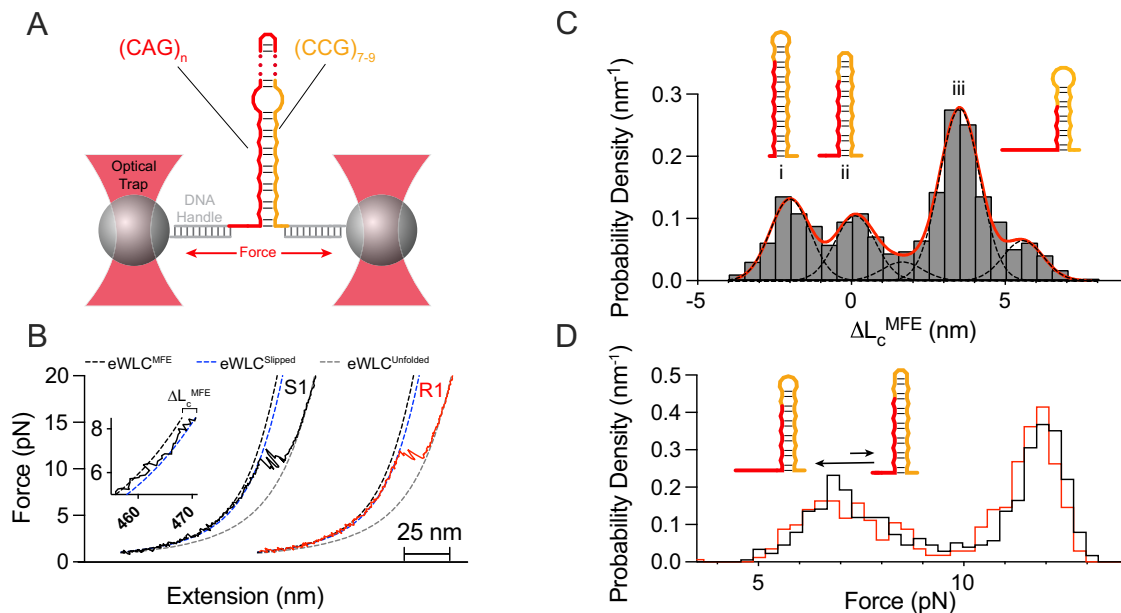
Fig. 1 | Normal HTT RNA with 7 CAG repeats unfolds cooperatively. A RNAs corresponding to the CAG repeat region of human HTT exon 1 were attached to optically trapped beads via dsDNA handles. fHTT$_n$ mRNAs contain $n = 7, 12, 20,$ or 40 CAG triplets (red line) plus flanking CCGCAA(CCG)$_7$ repeats (gold line). **B** Example stretch (black, S1) and relax (red, R1) FEC for fHTT$_7$ at 90 nm/s show unfolding/refolding transitions near 11.5 pN. Inset: slip transitions at low tension. Dashed curves represent eWLC models for folded (black), unfolded (gray) and slipped (blue) states. The expected ΔL_c for transitioning from the fully folded to the fully unfolded state is 26.0 nm. See Supplementary Fig. 4 for details. S1 and R1 are plotted with an offset for clarity. **C** Contour lengths of slipped intermediates before the main unfolding transition from eWLC fits (blue dashed line in (**B**)) relative to the folded state (eWLC^{MFE}—black dashed line in (**B**)). Red line, sum of 5 Gaussians; black dashed line, individual Gaussians. The histogram was fit using GraphPad Prism version 10.2.3 for Mac OS X. Labeled peak averages correlate with structures predicted by mfold[54] and footprinting[10]: (i) no unpaired CAG triplet (-1.76 ± 0.04 nm), (ii) reference MFE structure with one single-stranded CAG triplet (0.19 ± 0.06 nm), (iii) three unpaired CAG triplets (3.87 ± 0.02 nm). RNA schematics were prepared with RNAcanvas[55]. **D** Distributions of forces for unfolding (black) and refolding (red) transitions, from experiments as in (**B**). ($N = 286$ FECs from 48 molecules). The transitions represent slippage between intermediate structures (5–10 pN) and complete unfolding above 10 pN.

Why neurodegeneration rises around a threshold of 36 CAG repeats, and why some CAG expansion mutations are toxic whereas others are not, remain long-standing questions[21]. To address how repeat expansion changes the structural properties of *HTT* mRNA, we probed the folding and unfolding of *HTT* mRNA fragments using single-molecule force spectroscopy with optical tweezers that detect unfolding of a few base pairs at a time[22,23] and resolve heterogeneous folding pathways[24,25]. Our results show how continuous, non-cooperative slippage of CAG base pairs hampers complete refolding of exp*HTT* mRNA. "Sticky" interactions with the adjacent CCG trinucleotides stabilize partially slipped structures containing unpaired CAG repeats. The exposed single-strands readily initiate inter-strand base pairing, supporting models in which single-stranded CAG repeats drive phase separation and aggregation of CAG repeat RNA[18,26,27].

Results

Force spectroscopy resolves *HTT* RNA folding intermediates

To investigate *HTT* mRNA structures by force spectroscopy, fragments of human *HTT* exon 1 (f*HTT*) RNA spanning the CAG repeat region were joined with dsDNA handles and attached to polystyrene beads held by optical traps (Fig. 1A). The mRNA fragments contained 7–40 CAG codons, the adjacent CCG-CAA-(CCG)$_7$ codons, and an additional 20 nt *HTT* sequence flanking the trinucleotide repeats (Supplementary Fig. 1 and Supplementary Table 1). Tethered f*HTT* mRNAs were repeatedly stretched and relaxed at a constant speed, yielding force-extension curves (FECs) that revealed the size and mechanical stability of the mRNA structures (Fig. 1B). Fitting an extensible worm-like chain (eWLC) model to the data yielded the changes in contour length (L_c) upon unfolding and refolding (see "Methods" and Fig. 1B), the forces at which unfolding and refolding occur (see Methods and Supplementary Fig. 2) and the cooperativity of each transition (Supplementary Fig. 3).

mRNA containing 7 CAG triplets (f*HTT*$_7$), within the non-pathogenic range, exhibited near-equilibrium, cooperative unfolding and refolding transitions at 11.8 ± 0.6 pN (Fig. 1B). However, the distribution of contour length changes associated with these transitions ranged from 18.2 nm to 29.7 nm, compared to 26.0 nm expected for unfolding of the predicted minimal free energy (MFE) hairpin. This difference is explained by smaller transitions below ~7 pN (Fig. 1B inset), indicating that f*HTT*$_7$ mRNA populates "slipped" structures before unfolding completely at higher force.

To evaluate these slipped structures, we fitted each FEC with a series of eWLC models to determine the contour length change relative to L_c of the predicted MFE structure (ΔL_c^{MFE}) which is comprised of base paired CAG-CCG triplets plus one unpaired CAG (Supplementary Fig. 4). The distribution of ΔL_c^{MFE} at low tension peaked at values that differed by multiples of one single-stranded trinucleotide (~1.77 nm; Fig. 1C and Supplementary Table 2). This length change agreed well with "slipped" conformations of both shorter and longer contour lengths than the predicted MFE structure (Fig. 1C). Control experiments showed that these low force transitions did not arise from transitions in the *HTT* nucleotides flanking the repeat region (Supplementary Fig. 5). The FECs were not well described by sequential opening of the hairpin, indicating that the CAG/CCG repeat hairpin slips rather than partially unfolds at low forces.

The presence of slipped intermediates before the main unfolding transition was supported by two populations of the unfolding and refolding forces (Fig. 1D): transitions centered around 11.8 ± 0.6 pN were almost exclusively associated with cooperative unfolding of the f*HTT*$_7$ hairpin, whereas transitions at ~6.9 ± 1 pN corresponded to "slippage" of one to four CAG repeats.

Thus, an unexpanded *HTT* mRNA folds cooperatively into a hairpin that is stabilized by base pairs between CAG and CCG triplets. Yet,
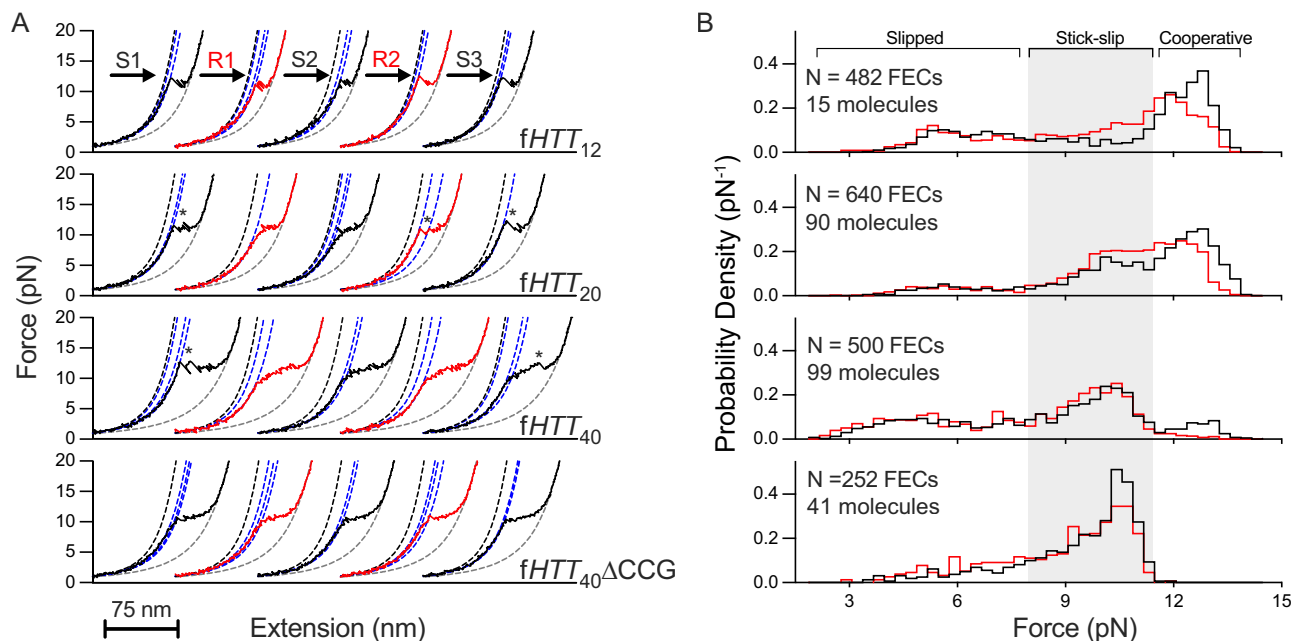
Fig. 2 | Expanded HTT mRNA unravels non-cooperatively. A Three successive cycles of stretching (black) and relaxation (red) for fHTT$_{12-40}$ mRNAs, fit to eWLC models as in Fig. 1. Cooperative transitions (asterisks) become less common as the repeat expands and are lost for ΔCCG (bottom). Expected ΔL$_c$ from the fully folded to the fully unfolded state is 36.58 nm for fHTT$_{12}$; 50.74 nm for fHTT$_{20}$; 86.14 nm for fHTT$_{40}$; 68.44 nm for fHTT$_{40}$ΔCCG. **B** Probability density of unfolding and refolding forces for stretching (black) and relaxation (red) for each fHTT mRNA. Average unfolding force for fHTT$_{12}$ was 12.4 pN, compared to 11.8 pN for fHTT$_7$. Shaded region indicates the force range for non-cooperative stick-slip unfolding transitions.

even this short mRNA samples different structures resulting from slippage of base pairs within the repeat region. These results agree well with previous single-molecule FRET studies of slippage in DNA hairpins containing pure CAG or CTG repeats[28–30], suggesting that such interconversion is an intrinsic property of nucleic acids containing triplet repeats.

Slippage leads to non-cooperative unfolding of expanded CAG repeats

Expanding the repeat region beyond seven CAG triplets resulted in progressive loss of large cooperative transitions during hairpin opening and closing (Fig. 2). fHTT$_{12}$, containing 12 CAG triplets, visited a greater number of slipped intermediates than fHTT$_7$, but still unfolded cooperatively at 12.3 ± 0.7 pN (Fig. 2A, first row). By contrast, expansion to 20 or 40 CAG triplets caused a marked reduction of folding cooperativity (Fig. 2A, second and third rows). Almost half of the FECs for fHTT$_{20}$ and nearly all FECs for fHTT$_{40}$ showed gradual non-cooperative extension between 8.5–11 pN that we attribute to stick slip unfolding of the repeat region (see below). Analogous non-cooperative unfolding was observed for a tetratricopeptide repeat protein[31] and nascent RNA[32]. As the mRNA structure unraveled, transitions to structures that were slightly more or less extended than the preceding conformation indicated reorganization of the base pairs that is coupled to progressive unfolding.

Due to hairpin slippage, non-cooperative unfolding of the repeat region occurred at lower forces than cooperative unfolding (8.5–11 pN vs. 12–12.5 pN, respectively; Fig. 2B), and the unfolding force remained constant as the CAG repeats expanded from 12 to 40. Initial and subsequent stretches of fHTT$_{40}$ populated similar slipped intermediates (Supplementary Fig. 6A). However, cooperative unfolding of fHTT$_{40}$ at high force was observed more frequently during the first stretch of a molecule (Fig. 2A and Supplementary Fig. 6B, C), indicating that the long CAG hairpin did not equilibrate at low tension. The fraction of cooperative transitions remained unchanged when the equilibration between force-ramp cycles was increased to 10 s (Supplementary

Fig. 6D), a time that is sufficient to equilibrate many RNA hairpins at room temperature[33–35].

Altogether, these results suggested that non-cooperative unraveling and contraction of the CAG region arises from frequent slippage of base pairs that becomes more prevalent as the repeats expand. Single-molecule FRET measurements showed that fHTT$_{40}$ RNA favors compact structures at zero force, yet a sizeable fraction of molecules adopt extended conformations (Supplementary Fig. 1E), as observed for CAG$_{20}$ RNA[36]. As the hairpins slip, exposed single-stranded bases pair again in a new register, reorganizing the structure (Fig. 3A). Applied force favors slippage and thus reorganization, explaining why we observe jumps to both shorter and longer structures as the tension rises (Fig. 3B). When relaxing the force, random nucleation of CAG hairpins leads to kinetically trapped, misaligned hairpins that unfold non-cooperatively during the subsequent stretch. Rapid opening and closing of CAG hairpins can account for the gradual extension (or contraction) of the RNA that we observe at 8.5–10 pN[32,37–39], although not reorganization of the repeat structure nor the small slip transitions at low force. A key feature of repeat RNAs is the sheer number of base pairing possibilities that favor compact structures (at low force) while lowering the barrier to conformational exchange.

Flanking CCG repeats introduce "sticky" interactions

We next asked what other features of fHTT mRNA contribute to its rugged folding landscape. A comparison of successive stretch and relaxation cycles of fHTT$_{40}$ mRNA showed that the unfolding path varied between pulls on the same molecule (Fig. 2A). Gradual extension of the repeat region was often punctuated by discrete unfolding transitions or rapid hopping between structures before the mRNA was fully unfolded (Fig. 2A, third row, last FEC). This variation contributed to the broad distribution of intermediates and unfolding forces.

As cooperative transitions occurred at higher forces than non-cooperative transitions (Fig. 2B), they must reflect the unfolding of mechanically stable structures. We conjectured that these stable structures arise from interactions with the flanking CCG repeats, which
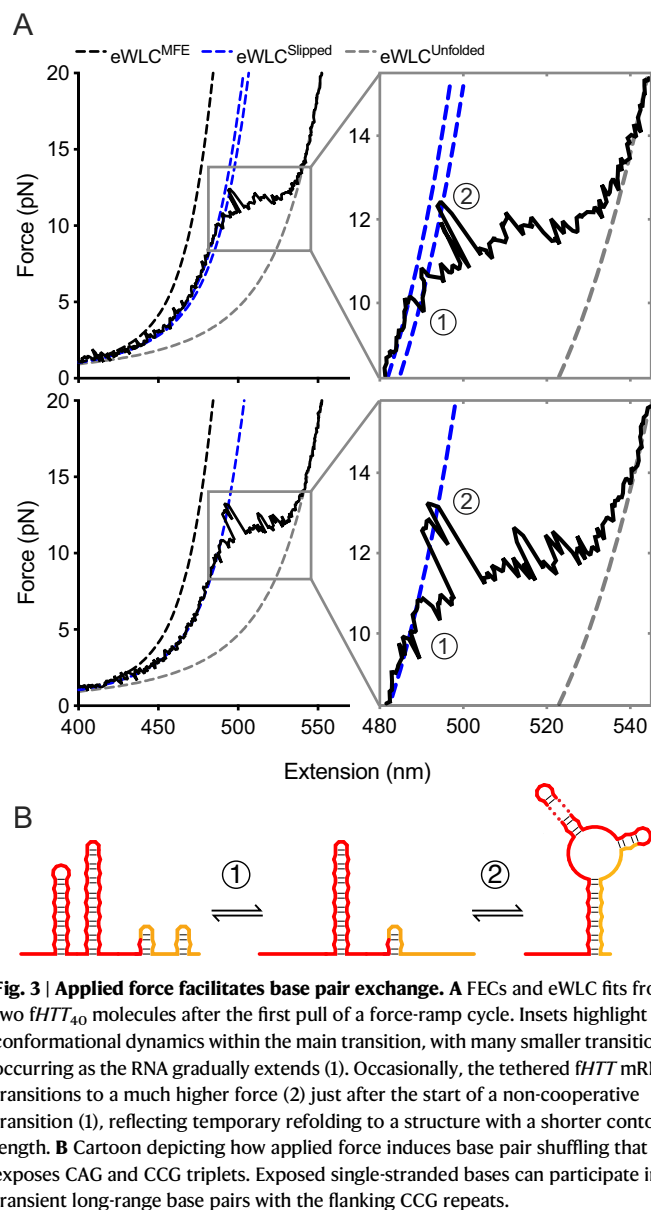
A

B

Fig. 3 | Applied force facilitates base pair exchange. A FECs and eWLC fits from two $fHTT_{40}$ molecules after the first pull of a force-ramp cycle. Insets highlight conformational dynamics within the main transition, with many smaller transitions occurring as the RNA gradually extends (1). Occasionally, the tethered $fHTT$ mRNA transitions to a much higher force (2) just after the start of a non-cooperative transition (1), reflecting temporary refolding to a structure with a shorter contour length. **B** Cartoon depicting how applied force induces base pair shuffling that exposes CAG and CCG triplets. Exposed single-stranded bases can participate in transient long-range base pairs with the flanking CCG repeats.

are known to increase the mRNA melting temperature[10]. Owing to sequence degeneracy, base pairing of the CCG repeats with any of the 40 CAG repeats can give rise to many intermediate structures.

To test this possibility, we measured the folding of $fHTT_{40}\Delta CCG$ mRNA containing only 40 CAG repeats but no CCG repeats. Deletion of the flanking CCG repeats eliminated the cooperative transitions, resulting in smooth non-cooperative unraveling (Fig. 2A, bottom). The elimination of cooperative transitions correlated with a narrower distribution of unfolding forces and fewer resolved slipped intermediates for $fHTT_{40}\Delta CCG$ than for $fHTT_{12}$, $fHTT_{20}$ or $fHTT_{40}$ (Fig. 2B, bottom).

We concluded that the cooperative transitions observed in $fHTT$ mRNAs correspond to folding and unfolding of hairpins with CAG-CCG base pairs at the base of the stem to which tension is being applied. Base pairing between CAG and CCG triplets at the ends of the repeat region becomes less likely as the number of CAG triplets increases, explaining the observed decline in large cooperative transitions (Fig. 2B). However, locally stable structures created by CCG repeats can unfold at any point during unraveling of the repeat region. Therefore, we propose that the more stable base pairs formed by the flanking CCG repeats create "sticky" interactions that hinder the facile slippage of CAG-CAG base pairs within expanded HTT mRNA.

Variable refolding in expanded *HTT* mRNAs

Inspection of individual FECs showed that $fHTT_{40}$ RNA traveled through many intermediate conformations at low force before unfolding at high force (Fig. 4A). To quantify the occupancy of slipped conformations, we calculated the contour lengths of all identifiable intermediates before the high force unfolding transition as in Fig. 1C (Methods). As anticipated, the distribution of intermediates with different ΔL_c^{MFE} values broadened as the number of CAG triplets expanded from 7 to 40 (Fig. 4B). Whereas $fHTT_7$ populated a few intermediates with up to four unpaired CAG triplets, $fHTT_{12}$ occupied a wider distribution of intermediates that shifted to even more extended conformations for $fHTT_{20}$ and $fHTT_{40}$ (Fig. 4B). These results are consistent with FRET experiments suggesting CAG20 RNA hairpins adopt a distribution of slipped, folded states[36]. Distinct peaks in the ΔL_c^{MFE} distribution separated by $\Delta L_c \sim 2$ CAG repeats suggested a preference for folding steps involving an even number of CAG triplets, similar to CAG DNA hairpins[28].

Deletion of the flanking CCG repeats narrowed the distribution of intermediates and shifted the population to shorter ΔL_c^{MFE} values, suggesting that pure CAG repeats refold into fewer distinguishable structures than $fHTT_{40}$ (Fig. 4B, bottom). This observation supported our conclusion that the CCG repeats impede relaxation of the mRNA structures and increase the number of structures formed.

Conformational entropy of exp*HTT* mRNA is inherent to its sequence

Our results indicated that the degeneracy of base pairing configurations hinder exp*HTT* mRNA from adopting any single structure. This interpretation was supported by the similar range of contour lengths sampled during the first and subsequent stretches of $fHTT_{40}$ mRNA (Supplementary Fig. 6A). These intermediates were not due to differences between molecules, as different $fHTT_{40}$ molecules sample similar distributions of L_c and F_U (Supplementary Fig. 7). Similar folding intermediates were observed in NaCl or KCl (Supplementary Fig. 8). Extended intermediates were more prevalent in Mg^{2+} or upon rapid release of tension ("snap relax"; Supplementary Fig. 8), presumably because partially folded structures are more easily trapped. This structural heterogeneity observed over a range of conditions suggested that high conformational entropy is an intrinsic feature of exp*HTT* mRNA.

Slippage occurs in sequential steps

To investigate how $fHTT_{40}$ mRNA dynamically slips between different secondary structures, we performed passive mode experiments in which the optical traps were kept at a constant separation[33], revealing slippage events with discrete changes in force over time (Fig. 4C, Methods). Changes in force were converted into ΔL_c^{MFE} and a hidden Markov model (HMM) was used to obtain the transition probabilities (Fig. 4C, Methods). At low force (~5–9 pN), the data were best described by ten states, with force-dependent dwell times ranging from 0.3 to 1.1 s (Supplementary Table 3). Each state differed by the contour length of one CAG trinucleotide (Supplementary Fig. 9). The transition density plot (TDP, Fig. 4D) revealed that most transitions occur between adjacent slipped conformations. These measurements established slippage by one trinucleotide as the fundamental unit of exp*HTT* conformational dynamics.

When $fHTT_{40}$ was held at a higher pre-tension near the global unfolding transition, hopping between conformations of similar ΔL_c^{MFE} was punctuated by brief excursions to more folded states (Supplementary Fig. 10, Supplementary Table 4). These unexpected transitions to shorter structures were associated with varied changes in ΔL_c^{MFE}. This variation can be explained by transient base pairing between the CCG repeats and different segments of the CAG tract that become exposed at high force. Thus, the force-ramp and passive mode results show that the expanded array of CAG repeats hops or slips
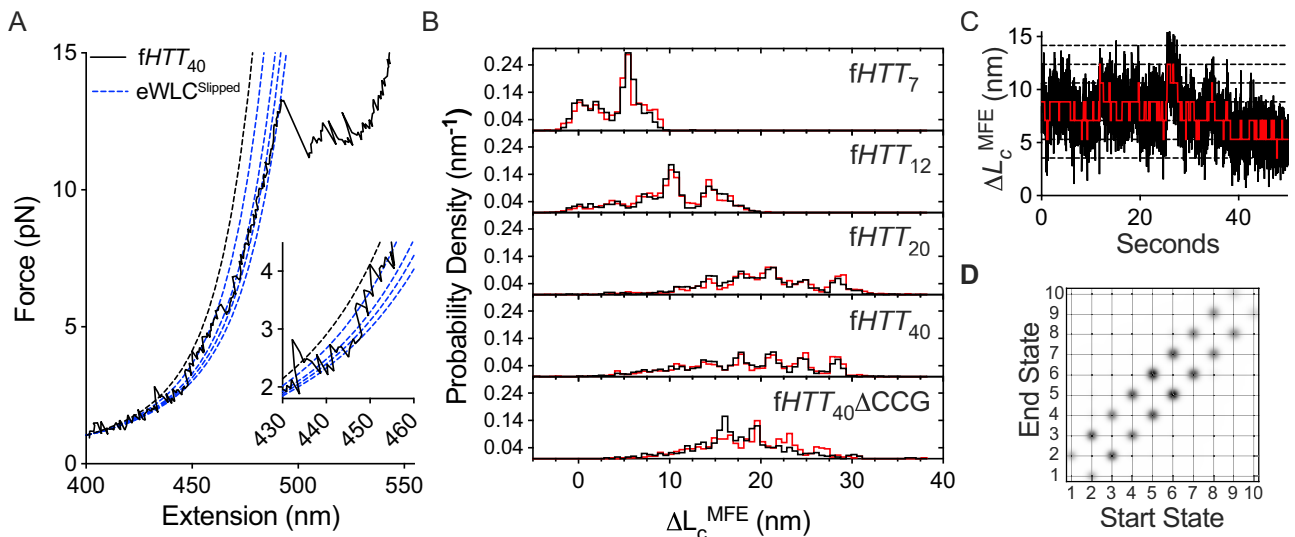
Fig. 4 | Continuous stepwise slippage of expanded CAG tract. Slipped intermediates of fHTT$_{40}$ analyzed as in Fig. 1. **A** Example FEC demonstrating transitions at low force that extend or shorten L$_c$, before unfolding at 13.2 pN; see Supplementary Figs. 2, 3 for analysis details. Black dashed line: eWLC model for fully folded fHTT$_{40}$. Blue dashed lines, eWLC fits to structures differing by 2–4 unpaired CAG triplets. Inset: Expansion of low force region. **B** Distribution of ΔL_c^{MFE} for all

fHTT mRNAs. Stretch, black; relax; red. **C** Passive mode fluctuations in ΔL_c^{MFE} for fHTT$_{40}$ at -6 pN pretension. Black dashed lines show HMM state, red line shows the HMM model. Similar results were obtained when the HMM fit was unconstrained (Supplementary Fig. 9). Hopping rates of 0.9–3 s^{-1} are comparable to the folding kinetics of a stable 21 bp hairpin[33,34]. D. Transition density plot from the HMM model in (**C**). See Supplementary Fig. 10 for data at high force.
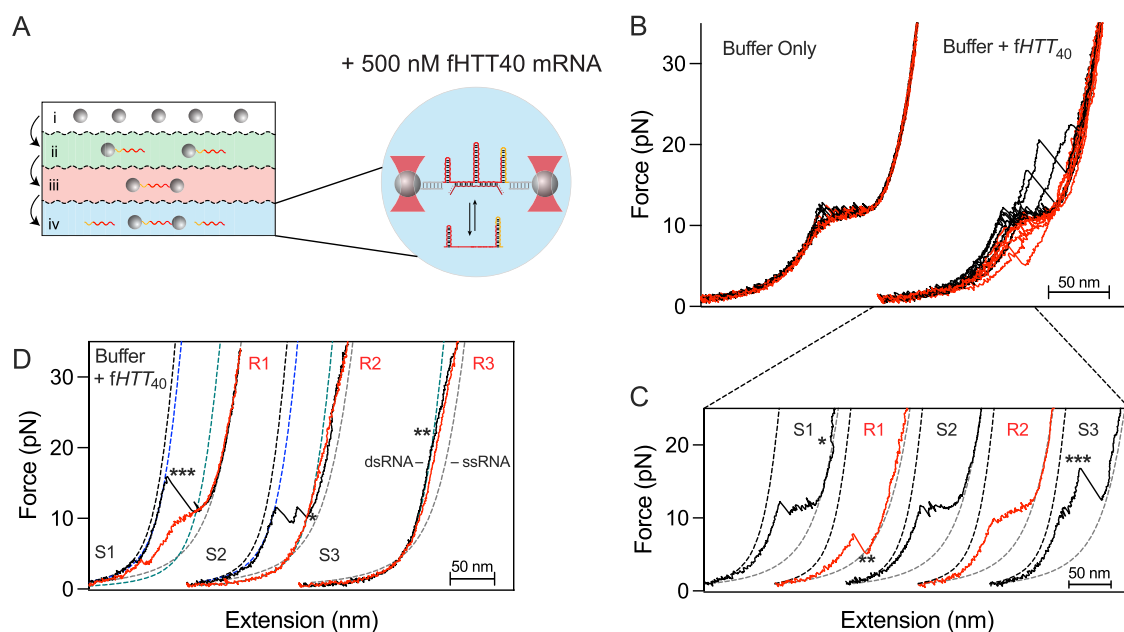

Fig. 5 | Intermolecular association of partially unfolded exp*HTT* RNA. A Laminar flow microfluidics for intermolecular interactions. fHTT$_{40}$ mRNA is tethered in buffer (channels i–iii) before exposure to 500 nM free fHTT$_{40}$ mRNA (channel iv), where it can form intermolecular base pairs. **B** Example stretch (black) and relax (red) FECs in buffer (left, channel iii) and for the same molecule after exposure to free mRNA (right, channel iv). **C** Consecutive cycles of stretching (S1–S3, black) and

relaxation (R1–R3, red) of a single fHTT$_{40}$ molecule, showing (*) free RNA binding at high force; (**) release at low force; (***) unfolding and partial release at high force. **D** Similar experiment as in (**C**) with another fHTT$_{40}$ molecule, showing conversion to dsRNA after (***) unfolding and full release of free mRNA; (*) binding and conversion to dsRNA during unfolding; (**) partial melting of the dsRNA at high force. Green dashed line, eWLC for double-stranded fHTT$_{40}$ mRNA.

between different base pairing registers that expose or sequester CAG trinucleotides.

Slipped CAG base pairs enable inter-strand pairing

We lastly asked whether the dynamic rearrangements observed for fHTT$_{40}$ could facilitate exp*HTT* mRNA aggregation. Unpaired CAG triplets exposed by non-cooperative slippage could serve as toeholds

for intermolecular base pairing, which is the first step of RNA aggregation[10,18]. To test this idea, we moved a single tethered fHTT$_{40}$ mRNA into a microfluidic channel containing 500 nM free fHTT$_{40}$ mRNA (Fig. 5A). Strikingly, we immediately observed new features in the FEC, compared to the preceding force-ramp cycles of the same tethered mRNA in buffer only (Fig. 5B). Because these features were never observed for isolated mRNA, they must be a direct result of
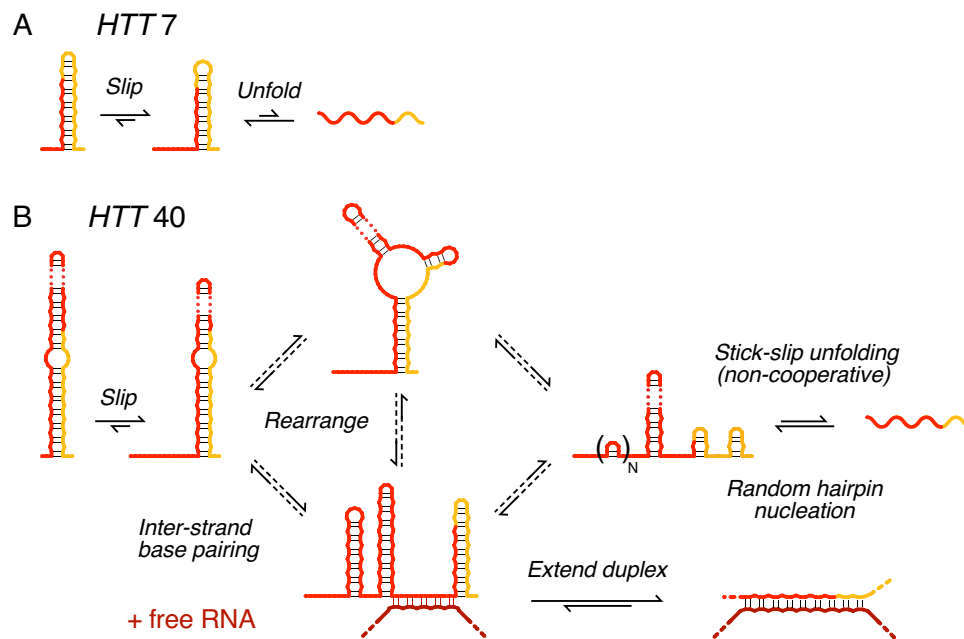
Fig. 6 | Stick-slip model for CAG repeat aggregation in HD. A Short CAG tracts form composite hairpins that unfold and refold cooperatively with limited slippage of base pairs. **B** Expanded CAG repeats slip and rearrange non-cooperatively, producing a mixture of secondary structures that include unstable, slippery CAG hairpins (red) and stable, sticky CAG•CCG hairpins that unfold in small rips (red and gold). Single-stranded regions nucleate base pairing with another RNA, producing a network that can lead to aggregation (bottom). Although transitions between specific intermediate structures are not resolved in our experiments, the cartoons illustrate conformations consistent with the force-extension results.

inter-strand interactions between tethered and freely diffusing $fHTT_{40}$.

In one example, the first extension in the presence of free mRNA was normal up to ~18 pN, at which point the unfolded tethered mRNA jumped to a shorter structure (Fig. 5C, *). Upon relaxation of the force, the mRNA did not refold gradually around 10 pN as usual, but instead refolded in a single step at 7.5 pN (Fig. 5C, **). These perturbations can be explained by partial inter-strand base pairing with a diffusing mRNA at high force that prevents the tethered mRNA from refolding, until dissociation of the diffusing mRNA at 5 pN allows the tethered $fHTT_{40}$ to reestablish intra-strand base pairs. In agreement with this interpretation, the second force-ramp cycle displayed the expected force-extension behavior (Fig. 5C, S2-R2).

In the third force-ramp cycle, the tethered $fHTT_{40}$ remained folded until abruptly opening at ~17 pN, far above its usual unfolding force (Fig. 5C, ***). Unfolding at high force can be explained by inter-strand base pairing that bridges the CAG tract like a cruciform, preventing the tethered RNA from unfolding until one side of the complex dissociates. Interactions between RNA molecules typically occurred near the unfolding force for the tethered RNA, indicating that partial unfolding of the tethered RNA favors inter-strand base pairing. Exposed CAG triplets in the free RNA likely also contribute.

After a few force-ramp cycles in the presence of free RNA, several $fHTT_{40}$ molecules failed to refold, and remained extended during subsequent stretch-relax cycles (Fig. 5D, **). The FECs for these molecules were described well by the eWLC of an RNA duplex the length of $fHTT_{40}$, with a persistence length of 40 nm (Fig. 5D, green dashed line). Therefore, inter-strand base pairing, once initiated, can persist, and ultimately replace the self-structure of the tethered RNA.

Discussion
Our single-molecule force spectroscopy results reveal an unanticipated dynamical transition in expHTT mRNA as the CAG repeat number increases (Fig. 2). Short $fHTT$ mRNAs with 7 or 12 CAG repeats unfold and refold cooperatively, indicating that they form defined

hairpins, as previously proposed[10] (Fig. 6). By contrast, RNAs with 40 CAG repeats almost exclusively exhibit stick slip unfolding. This shift in behavior suggests that the absence of stable, cooperatively unfolding structures contributes to pathogenicity that is observed above a threshold of 36 CAG repeats[1,40]. This accordion-like unraveling and refolding of the expanded CAG tract arises from constant slippage and reorganization of CAG base pairs, and from heterogeneous nucleation of new hairpins (Fig. 6)[37]. As a result, the expanded repeats rarely refold completely. Importantly, the CCG repeats adjacent to the CAG tract in HTT form stable base pairs with many segments of the expanded CAG tract, kinetically trapping the RNA in a variety of structures. We show that stick-slip unfolding arising from a mix of CAG-CAG and CAG-CCG base pairs in expHTT mRNA favors extended structures (Fig. 4) that readily interact with other expHTT mRNA molecules once the repeat tract expands to 20–40 CAG triplets (Fig. 5).

Although we observe slippage at low tension, reorganization of the CAG tract is stimulated by applied force (Fig. 2 and Fig. 4), consistent with destabilization of one or more base pairs during the transitions between slipped conformations. Slip transitions could arise from end fraying[26] followed by rapid formation of new base pairs, or the propagation of internal bubbles, as recently proposed for DNA repeats[30]. Alternatively, the two strands of a CAG hairpin may slide past each other via continuous exchange of hydrogen bonds between neighboring nucleobases. This type of exchange does not interrupt base stacking and can occur quickly in double-stranded RNA[41] and DNA homopolymers[42].

Our experiments do not resolve internal rearrangements of the structures that contribute to the observed slip transitions. Such rearrangements are more favorable for repetitive sequences that can base pair in many similar configurations, however, than for heterogeneous sequences that have limited ways of refolding. An interesting question is why the expanded CAG repeats unravel around 10 pN. One answer is that base pair shuffling at low tension maintains the average amount of base pairing, in agreement with the short end-to-end distance of $fHTT_{40}$ RNA in solution (Supplementary Fig. 1E)[36]; Once the applied

tension exceeds the threshold for stick-slip unfolding, CAG hairpins are more likely to open whereas transitions to shorter structures become less probable, and the RNA starts to extend. Stick-slip unfolding occurs at lower forces than cooperative unfolding because it involves smaller transitions or reorganization of weaker CAG-CAG base pairs.

It has been recently suggested that the capacity of repeat RNAs to form heterogeneous interactions drives phase separation into droplets or aggregates[18,27,43] that have been linked to neurotoxicity[6,10,44]. However, temperature-jump experiments show that CAG20 RNA hairpins remain folded in cells, suggesting that partial unfolding of the CAG tract is needed to nucleate interactions between RNAs[45], consistent with the positive effect of thermal renaturation on phase separation in vitro[11]. We observe that the mixed CAG/CCG repeats present in *HTT* mRNA produce a "stick-slip" behavior that impedes base pair exchange, limiting the number of structures sampled by the RNA. The CCG codons may reinforce the folded structures of short *HTT* RNAs and inhibit their condensation relative to pure CAG repeats, while stabilizing interactions between expanded *HTT* RNAs. We note that HD is one of at least two trinucleotide repeat expansion disorders in which flanking sequences interact with the CAG repeats[27]. It is not known whether these mixed sequences influence aggregation of longer RNAs in cells or somatic expansion of the repeat locus.

Since unraveling exp*HTT* mRNA (10.0 ± 0.7 pN) requires less force than cooperatively unfolding shorter hairpins, our results also explain why ribosomes that exert ~13 pN on the mRNA[46] readily pass through expanded CAG tracts. Transient unfolding by ribosomes or helicases is likely to increase the heterogeneity of the RNA structure because long-range base pairs will not have sufficient time to reequilibrate[47]. However, transient unfolding is unlikely to dissolve repeat RNA condensates because the network of RNA interactions remodels without complete strand dissociation, as observed in our experiments. In cells, the RNA dynamics is likely influenced by RNA-binding proteins and helicases that locally perturb the CAG base pairs. The single-molecule platform described here will be useful for probing the effects of proteins or small molecules on folding and aggregation of repeat-containing RNA.

Methods
Sample preparation
Double-stranded (ds) DNA handles with single-stranded overhangs were prepared via PCR amplification of pROEX_HTa[48]. Primers used to generate the dsDNA handles are listed in Supplementary Table 1. A 691 bp biotinylated DNA handle with a 29 nt 5′ overhang was amplified by Q5 DNA polymerase (NEB, Cat# M0491S) using a 5′ biotinylated forward primer and a reverse primer containing an abasic site[49]. A 673 bp digoxigenin-labeled DNA handle was amplified in Q5U PCR master mix (NEB, Cat# M0597L) with a forward primer containing a deoxyuridine and a reverse primer with a 5′ digoxigenin tag. A 30 nt 3′ overhang was created by treating the purified PCR product with USER enzyme mix (NEB, Cat# M5505S) for 15 min at 37 °C. All PCR products were cleaned up (Qiagen, Cat# 28104) prior to use.

Transcription templates for fragments of the human *huntingtin* exon 1 (f*HTT*N) were synthesized and cloned into pUC-GW-Amp (Genewiz) to yield pUC-GW-f*HTT*N, in which N designates the number of CAG triplets. Transcription templates contained a T7 promoter, 29 bp complementary to the 5′ overhang of the biotinylated DNA handle, 18 bp native exon 1 sequence, N CAG triplets, CCGCAA(CCG)$_7$ (flanking polymorphic repeats), a further 20 bp of native exon 1 sequence downstream of the repeats, 30 bp complementary to the digoxigenin-labeled DNA handle and a T7 terminator. Predicted secondary structures of f*HTT* mRNA fragments are shown in Supplementary Fig. 1. Transcription templates were prepared by PCR using primers complementary to the T7 promoter and terminator sequences (see Supplementary Table 1) or by digestion of the plasmid DNA with

SphI (NEB, Cat# R3182S). mRNA fragments were prepared by in vitro transcription with T7 RNA polymerase using a ratio of 10:30 AU:GC NTPs. The desired transcripts were purified on denaturing 6−8% polyacrylamide gels.

Purified f*HTT* mRNA transcripts were combined with the dsDNA handles at 25 nM equimolar ratio, and annealed and refolded by denaturation at 85 °C for 10 min, 65 °C for 1.5 h, 55 °C for 1.5 h and slow cooling to 4 °C (0.1 °C/s) in 200 mM NaCl, 20 mM PIPES, pH 6.5, 60% formamide[50]. Formation of the desired DNA-RNA complexes were verified by agarose gel electrophoresis (Supplementary Fig. 1), purified using a PCR clean up kit (NEB, Cat# T1030L) and eluted in water. Freshly prepared refolded/annealed complexes (1 μL 25 nM stock; ~2.5 nM final) were combined with 4 μL 0.1% w/v 2.1 μm diameter polystyrene beads coated with anti-digoxigenin antibodies (Spherotech, Cat# DIGP-20-2), diluted to 10 μL with assay buffer (10 mM MOPS, pH 7, 250 mM NaCl, 1 mM EDTA), and incubated for 10 min at room temperature. The bead mixture was then diluted ~50 fold in assay buffer spiked with RNase inhibitor. 1 μL 0.5% w/v streptavidin coated 2.1 μm diameter polystyrene beads (Spherotech, Cat# SVP-20-5) were diluted into 1 mL assay buffer. The two bead mixtures were loaded into separate channels of the Lumicks C-Trap microfluidics system.

Force-ramp data collection
Force-ramp assays were performed using a Lumicks C-Trap instrument with BlueLake software (Lumicks). Briefly, streptavidin and anti-digoxigenin coated beads were collected by two optical traps and brought to a separate channel containing assay buffer spiked with RNase inhibitor, 10 mM sodium azide and protocatechuic acid/protocatechuate-3,4-dioxygenase[51]; for oxygen scavenging. The optical traps were calibrated to obtain the trap stiffness for each pair of beads. Single molecules were tethered between two beads by slowly bringing the trapped beads together until a small increase in force was observed. Once a tether was made, the optical traps were separated at a constant velocity of 90 nm/s to unfold the mRNA and brought together at the same speed to refold the mRNA. Single tethers exhibited the expected persistence length for the DNA handles and folding transitions below 20 pN. At this pulling speed, the loading rate was estimated to be ~10 pN/s just before the start of the main transition of the force vs time traces. After relaxation, the trap movement was paused for 1 s before repeating the stretch and relaxation cycle. For each f*HTT* mRNA, 20 or more molecules were tethered, and 1-100 force-ramp cycles were performed on each tether for each experiment.

Force-ramp data analysis and statistics
For each bead pair, a force baseline was collected and subtracted from valid FECs. Data were collected at 75 kHz and down-sampled to 100 Hz for analysis, and the molecular extension for each tether was calculated using the Lumicks Pylake Python package. In our dual trap instrument (Lumicks C-Trap), one of the traps (trap 1) is steered to apply force to a tethered molecule, whereas the other (trap 2) is static. Force is measured by each trap independently. FECs were calculated using the force recorded by trap 1.

Processed FECs were fit by the extensible worm-like chain (eWLC) model

$$F(x) = \frac{k_{\mathrm{B}}T}{L_{\mathrm{p}}} \left[\frac{1}{4}\left(1 - \frac{x}{L_{\mathrm{c}}} + \frac{F}{K}\right)^{-2} - \frac{1}{4} + \frac{x}{L_{\mathrm{c}}} - \frac{F}{K} \right] \qquad (1)$$

in which L_{c} is the contour length, L_{p} is the persistence length and K is the stretch modulus[52]. Extension of the folded and unfolded states were modeled by a series of eWLCs describing extension of the dsDNA handles, the ssRNA flanking the CAG hairpin(s) in the folded mRNA and opening of the mRNA secondary structure. L_{c} of the dsDNA handles was fixed at 473.6 nm based on the crystallographic contour length of dsDNA at 0.34 nm/bp, and L_{p} and K were fixed at

40 nm and 1000 pN, respectively, from reported values[23]. For the ssRNA segments, we used $L_p = 1$ nm and $K = 1500$ pN, leaving the L_c^{ssRNA} as the only free fitted parameter[23]. Extension of the unfolded *fHTT* mRNA was modeled by complete opening of the predicted hairpin structure at 0.59 nm/nt. The unfolded region of each FEC was aligned horizontally (x) to the high force region (15–20 pN) of the WLC model of the fully unfolded *fHTT* mRNA, followed by a vertical alignment (F) to the low force region of FECs and the WLC for the handles + folded *fHTT* mRNA as previously described[24]. All analyses were performed on aligned FECs.

Contour length. The ΔL_c associated with complete opening of the mRNA structure at high force was determined by fitting the region of the FEC just before and after the transition to two eWLC models in series, one for the dsDNA handles and another to describe the extension of the *HTT* ssRNA flanking the folded hairpin. Small changes in contour length below 9 pN were quantified by fitting intermediate regions of FECs that persisted for at least 30 ms with two eWLCs in series for the dsDNA handles and ssRNA regions, adapted from Ref. 38. The unfolded ΔL_c^{MFE} of each intermediate conformation was obtained from the fitted L_c value of the ssRNA segment in that conformer, after subtracting the L_c of the reference MFE structure (see Supplementary Fig. 4 for a detailed description).

To determine whether the single-stranded *HTT* exon 1 nucleotides flanking the CAG region contributed to the observed low force transitions, the overhangs of the dsDNA handles were extended to hybridize with these nucleotides. Force ramp experiments using these extended dsDNA handles showed similar results compared to the dsDNA handles with shorter overhangs (Supplementary Fig. 5). A control experiment using an unstructured DNA oligonucleotide to bridge the dsDNA handles, leaving 25 ssDNA nucleotides between the handles, indicated that the observed intermediate transitions at low force cannot be explained by fluctuations of the dsDNA handles (Supplementary Fig. 11).

Force. The force associated with each intermediate transition in each FEC was determined using a custom script (Python 3.9 and later) that determines the point at which the observed FEC deviates consistently from its fitted eWLC for the folded mRNA. Briefly, all identifiable intermediates in the folded region of each FEC were fit as described above, and the root-mean-squared error (RMSE) of the fitted eWLC from the experimental FEC was calculated using a sliding window. The unfolding force was defined as the first point at which the RMSE permanently crossed a threshold set to the mean + 3σ of the RMSE before the transition (Supplementary Fig. 2). All transitions (cooperative and non-cooperative) were analyzed in the same manner. Unfolding and refolding forces for transitions between slipped intermediates and between intermediate structures and the unfolded RNA were determined in the same way.

Cooperativity. The cooperativity of the main unfolding transition for a subset of FECs of $fHTT_{40}$ was defined by the deviation in dF/dx following the start of the transition. <dF/dx> and the associated standard deviation (σ) were defined over a slice of the FEC used for fitting the eWLC up to the start of the transition. Then, dF/dx was evaluated over a window starting from the fitted slice and extending 5 nm past the transition. Transitions were classified as cooperative when dF/dx in this region dropped below a threshold equal to <dF/dx> − 3.5σ (Supplementary Fig. 3A). Transitions were classified as non-cooperative when dF/dx failed to drop below this threshold within this region (Supplementary Fig. 3B).

Passive mode assays and HMM fitting with statistical analysis

After collecting FECs for a single tethered molecule, a pretension was set at low (~5 pN) and high (~12 pN) forces corresponding to the transitions observed in the FEC. Fluctuations in the force and extension at constant trap separation ("passive mode", PM) were recorded for ~5 min at each pretension before manually relaxing and stretching the tethered RNA. Single trajectories of force versus time, $F(t)$, were converted to unfolded contour length relative to the MFE structure, $\Delta L_c^{MFE}(t)$, using the eWLC equation described above and the parameter_trace function from the Lumicks Pylake package.

For the low force region of the FECs corresponding to slipped transitions, seven PM traces were collected from five different molecules with pretensions between 4.5 and 9 pN. Following conversion of the data into $\Delta L_c^{MFE}(t)$, traces were concatenated and fit using a step-finding HMM algorithm[53]. Processed data were first fit to a maximum likelihood step finding algorithm. The data were clustered using a Gaussian Mixture Model (GMM) and the number of states was increased by one until the Bayesian information criterion of the model reached a minimum[45]. The results of step finding and GMM clustering were then used for parameter initialization and fitting of the HMM model. Data were initially fit to a blind HMM to determine the number of states and their positions, yielding a model with 10 states whose distributions of dwell times can be described well by a single exponential function (Supplementary Fig. 9). A second model in which the separation between states was set to one CAG triplet (1.77 nm) agreed well with the blind HMM (Fig. 4C, D, Supplementary Table 3). PM traces at higher force, near the unfolding transition, were collected from four molecules at a pretension of ~11.8 pN and fitted as described above (Supplementary Fig. 10). The dwell time distributions for each state were fitted to a single exponential function by maximum likelihood estimation to extract the lifetime using the "dwelltimes" module from the Lumicks Pylake package. State means and lifetimes of all models are reported in Supplementary Tables 3 and 4.

Reporting summary

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

Data availability

The data reported in this paper and information required to reanalyze the data are available from the Johns Hopkins Research Data Repository at https://doi.org/10.7281/T1/GTUUUG.

Code availability

Custom python scripts useful for reanalyzing the data in this study are available from the Johns Hopkins Research Data Repository at https://doi.org/10.7281/T1/GTUUUG.

References

1. Paulson, H. Repeat expansion diseases. *Handb. Clin. Neurol.* **147**, 105–123 (2018).
2. Jarosińska, O. D. & Rüdiger, S. G. D. Molecular strategies to target protein aggregation in Huntington's disease. *Front. Mol. Biosci.* **8**, 769184 (2021).
3. Marti, E. R. N. A toxicity induced by expanded CAG repeats in Huntington's disease. *Brain Pathol.* **26**, 779–786 (2016).
4. Heinz, A., Nabariya, D. K. & Krauss, S. Huntingtin and Its role in mechanisms of RNA-mediated toxicity. *Toxins* **13**, 487 (2021).
5. Schwartz, J. L., Jones, K. L. & Yeo, G. W. Repeat RNA expansion disorders of the nervous system: post-transcriptional mechanisms and therapeutic strategies. *Crit. Rev. Biochem Mol. Biol.* **56**, 31–53 (2021).
6. Sun, X. et al. Nuclear retention of full-length *HTT* RNA is mediated by splicing factors MBNL1 and U2AF65. *Sci. Rep.* **5**, 12521 (2015).
7. Genetic Modifiers of Huntington's Disease (GeM-HD) Consortium. CAG repeat not polyglutamine length determines timing of Huntington's disease onset. *Cell* **178**, 887–900.e814 (2019).
8. Wright, G. E. B. et al. Length of uninterrupted CAG, Independent of polyglutamine size, results in increased somatic instability, hastening onset of Huntington disease. *Am. J. Hum. Genet.* **104**, 1116–1126 (2019).

9. Yuan, Y. et al. Muscleblind-like 1 interacts with RNA hairpins in splicing target and pathogenic RNAs. *Nucleic Acids Res.* **35**, 5474–5486 (2007).

10. de Mezer, M., Wojciechowska, M., Napierala, M., Sobczak, K. & Krzyzosiak, W. J. Mutant CAG repeats of Huntingtin transcript fold into hairpins, form nuclear foci and are targets for RNA interference. *Nucleic Acids Res.* **39**, 3852–3863 (2011).

11. Jain, A. & Vale, R. D. RNA phase transitions in repeat expansion disorders. *Nature* **546**, 243–247 (2017).

12. Morriss, G. R. & Cooper, T. A. Protein sequestration as a normal function of long noncoding RNAs and a pathogenic mechanism of RNAs containing nucleotide repeat expansions. *Hum. Genet.* **136**, 1247–1263 (2017).

13. Sznajder, L. J. & Swanson, M. S. Short tandem repeat expansions and RNA-mediated pathogenesis in myotonic dystrophy. *Int J. Mol. Sci.* **20**, 3365 (2019).

14. Nalavade, R., Griesche, N., Ryan, D. P., Hildebrand, S. & Krauss, S. Mechanisms of RNA-induced toxicity in CAG repeat disorders. *Cell Death Dis.* **4**, e752 (2013).

15. Rudnicki, D. D. et al. Huntington's disease–like 2 is associated with CUG repeat-containing RNA foci. *Ann. Neurol.* **61**, 272–282 (2007).

16. Sobczak, K., de Mezer, M., Michlewski, G., Krol, J. & Krzyzosiak, WJ. RNA structure of trinucleotide repeats associated with human neurological diseases. *Nucleic Acids Res.* **31**, 5469–5482 (2003).

17. Kiliszek, A., Kierzek, R., Krzyzosiak, W. J. & Rypniewski, W. Atomic resolution structure of CAG RNA repeats: structural insights and implications for the trinucleotide repeat expansion diseases. *Nucleic Acids Res.* **38**, 8370–8376 (2010).

18. Nguyen, H. T., Hori, N. & Thirumalai, D. Condensates in RNA repeat sequences are heterogeneously organized and exhibit reptation dynamics. *Nat. Chem.* **14**, 775–785 (2022).

19. Tauber, D. et al. Modulation of RNA condensation by the DEAD-box protein eIF4A. *Cell* **180**, 411–426.e416 (2020).

20. Busan, S. & Weeks, K. M. Role of context in RNA structure: flanking sequences reconfigure CAG motif folding in huntingtin exon 1 transcripts. *Biochemistry* **52**, 8219–8225 (2013).

21. Lee, D. Y. & McMurray, C. T. Trinucleotide expansion in disease: why is there a length threshold? *Curr. Opin. Genet. Dev.* **26**, 131–140 (2014).

22. Tinoco, I. Jr, Li, P. T. & Bustamante, C. Determination of thermodynamics and kinetics of RNA reactions by force. *Q Rev. Biophys.* **39**, 325–360 (2006).

23. Greenleaf, W. J., Woodside, M. T. & Block, S. M. High-resolution, single-molecule measurements of biomolecular motion. *Annu Rev. Biophys. Biomol. Struct.* **36**, 171–190 (2007).

24. Neupane, K. et al. Structural dynamics of single SARS-CoV-2 pseudoknot molecules reveal topologically distinct conformers. *Nat. Commun.* **12**, 4749 (2021).

25. Rico-Pasto, M., Alemany, A. & Ritort, F. Force-dependent folding kinetics of single molecules with multiple intermediates and pathways. *J. Phys. Chem. Lett.* **13**, 1025–1032 (2022).

26. Maity, H., Nguyen, H. T., Hori, N. & Thirumalai, D. Odd-even disparity in the population of slipped hairpins in RNA repeat sequences with implications for phase separation. *Proc. Natl Acad. Sci. USA* **120**, e2301409120 (2023).

27. Krzyzosiak, W. J. et al. Triplet repeat RNA structure and its role as pathogenic agent and therapeutic target. *Nucleic Acids Res.* **40**, 11–26 (2012).

28. Xu, P., Pan, F., Roland, C., Sagui, C. & Weninger, K. Dynamics of strand slippage in DNA hairpins formed by CAG repeats: roles of sequence parity and trinucleotide interrupts. *Nucleic Acids Res.* **48**, 2232–2245 (2020).

29. Hu, T., Morten, M. J. & Magennis, S. W. Conformational and migrational dynamics of slipped-strand DNA three-way junctions containing trinucleotide repeats. *Nat. Commun.* **12**, 204 (2021).

30. Ni, C. W., Wei, Y. J., Shen, Y. I. & Lee, I. R. Long-range hairpin slippage reconfiguration dynamics in trinucleotide repeat sequences. *J. Phys. Chem. Lett.* **10**, 3985–3990 (2019).

31. Synakewicz, M. et al. Unraveling the mechanics of a repeat-protein nanospring: from folding of individual repeats to fluctuations of the superhelix. *ACS Nano* **16**, 3895–3905 (2022).

32. Dalal, R. V. et al. Pulling on the nascent RNA during transcription does not alter kinetics of elongation or ubiquitous pausing. *Mol. Cell* **23**, 231–239 (2006).

33. Liphardt, J., Onoa, B., Smith, S. B., Tinoco, I. J. & Bustamante, C. Reversible unfolding of single RNA molecules by mechanical force. *Science* **292**, 733–737 (2001).

34. Wen, J. D. et al. Force unfolding kinetics of RNA using optical tweezers. I. Effects of experimental variables on measured results. *Biophys. J.* **92**, 2996–3009 (2007).

35. Woodside, M. T. et al. Direct measurement of the full, sequence-dependent folding landscape of a nucleic acid. *Science* **314**, 1001–1004 (2006).

36. Hautke, A. C. & Ebbinghaus, S. Folding stability and self-association of a triplet-repeat (CAG) RNA hairpin in cytomimetic media. *ChemSystemsChem* **3**, e2000052 (2021).

37. Rissone, P., Bizarro, C. V. & Ritort, F. Stem-loop formation drives RNA folding in mechanical unzipping experiments. *Proc. Natl Acad. Sci. USA* **119**, e2025575119 (2022).

38. Sen Mojumdar, S. et al. Partially native intermediates mediate misfolding of SOD1 in single-molecule folding trajectories. *Nat. Commun.* **8**, 1881 (2017).

39. Jonchhe, S. et al. Ensemble sensing using single-molecule DNA copolymers. *Anal. Chem.* **92**, 13126–13133 (2020).

40. Ciesiolka, A., Jazurek, M., Drazkowska, K. & Krzyzosiak, W. J. Structural characteristics of simple RNA repeats associated with disease and their deleterious protein interactions. *Front. Cell Neurosci.* **11**, 97 (2017).

41. Han, G. & Xue, Y. Rational design of hairpin RNA excited states reveals multi-step transitions. *Nat. Commun.* **13**, 1523 (2022).

42. Woodson, S. A. & Crothers, D. M. Preferential location of bulged guanosine internal to a G.C tract by 1H NMR. *Biochemistry* **27**, 436–445 (1988).

43. Wadsworth, G. M. et al. RNAs undergo phase transitions with lower critical solution temperatures. *Nat. Chem.* **15**, 1693–1704 (2023).

44. Li, L. B., Yu, Z., Teng, X. & Bonini, N. M. RNA toxicity is a component of ataxin-3 degeneration in Drosophila. *Nature* **453**, 1107–1111 (2008).

45. Hautke, A. et al. CAG-repeat RNA hairpin folding and recruitment to nuclear speckles with a pivotal role of ATP as a cosolute. *J. Am. Chem. Soc.* **145**, 9571–9583 (2023).

46. Liu, T. et al. Direct measurement of the mechanical work during translocation by the ribosome. *Elife* **3**, e03406 (2014).

47. Małecka, E. M. & Woodson, S. A. Ribosomes clear the way for siRNA targeting. *Nat. Struct. Mol. Biol.* **27**, 775–777 (2020).

48. Maciuba, K., Zhang, F. & Kaiser, C. M. Facile tethering of stable and unstable proteins for optical tweezers experiments. *Biophys. J.* **120**, 2691–2700 (2021).

49. Gál, J., Schnell, R., Szekeres, S. & Kálmán, M. Directional cloning of native PCR products with preformed sticky ends (autosticky PCR). *Mol. Gen. Genet.* **260**, 569–573 (1999).

50. Stephenson, W., Wan, G., Tenenbaum, S. A., Li, P. T. Nanomanipulation of single RNA molecules by optical tweezers. *J. Vis. Exp.* **(90)**, 51542 (2014).

51. Aitken, C. E., Marshall, R. A. & Puglisi, J. D. An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments. *Biophys. J.* **94**, 1826–1835 (2008).

52. Marko, J. F. & Siggia, E. D. Stretching DNA. *Macromolecules* **28**, 8759–8770 (1995).

53. Liu, H. & Shima, T. A fast and objective hidden Markov modeling for accurate analysis of biophysical data with numerous states. *bioRxiv*, (2021). 2021.2005.2030.446337.
54. Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**, 3406–3415 (2003).
55. Johnson, P. Z. & Simon, A. E. RNAcanvas: interactive drawing and exploration of nucleic acid structures. *Nucleic Acids Res.* **51**, W501–W508 (2023).

Author contributions
B.M.O. designed and performed experiments, analyzed and interpreted data, wrote and edited the paper. R.M. designed and performed initial experiments and edited the paper. G.J.A. wrote scripts, analyzed data and edited the paper. N.R. performed experiments and edited the paper. C.M.K. conceived experiments, interpreted data, edited the paper and obtained funding. S.A.W. conceived experiments, interpreted data, edited the paper and obtained funding.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-52764-x.

Correspondence and requests for materials should be addressed to Christian M. Kaiser or Sarah A. Woodson.

Peer review information *Nature Communications* thanks Simon Ebbinghaus, Alexander Hautke and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions 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.