MOLECULAR BIOLOGY

Massively parallel analysis of single-molecule dynamics on next-generation sequencing chips

J. Aguirre Rivera[1]†, G. Mao[1]†, A. Sabantsev[1]†, M. Panfilov[1]†, Q. Hou[1], M. Lindell[2], C. Chanez[3], F. Ritort[4,5], M. Jinek[3], S. Deindl[1]*

Single-molecule techniques are ideally poised to characterize complex dynamics but are typically limited to investigating a small number of different samples. However, a large sequence or chemical space often needs to be explored to derive a comprehensive understanding of complex biological processes. Here we describe multiplexed single-molecule characterization at the library scale (MUSCLE), a method that combines single-molecule fluorescence microscopy with next-generation sequencing to enable highly multiplexed observations of complex dynamics. We comprehensively profiled the sequence dependence of DNA hairpin properties and Cas9-induced target DNA unwinding-rewinding dynamics. The ability to explore a large sequence space for Cas9 allowed us to identify a number of target sequences with unexpected behaviors. We envision that MUSCLE will enable the mechanistic exploration of many fundamental biological processes.

Fundamental biological processes involve macromolecules that often exhibit complex dynamics that are intimately tied to function. Single-molecule techniques are ideally poised to characterize such complex dynamic processes because they avoid ensemble averaging that can obscure kinetic intermediate states and alternative kinetic pathways (1–18).

However, existing single-molecule approaches are typically limited to the study of a small number of different samples. At the same time, a large sequence or chemical space often must be explored to derive a comprehensive understanding of complex biological processes. Emerging ensemble methods for multiplexed in vitro investigations have profoundly influenced mechanistic biology (19–36). Most of these approaches use DNA sequencing to differentiate between library members. The DNA sequence can then be used as a direct readout or report on RNA or protein sequence by means of the central dogma or DNA barcoding. DNA barcoding, more generally speaking, extends these approaches to any type of chemical diversity (22, 24). Despite their great utility, all of these approaches have limited capacity to resolve complex dynamics owing to ensemble averaging. On the single-molecule side, data acquisition was recently reported on mixtures of two (37) or nine (38) different DNA sequences identified using fluorophore-labeled detection oligonucleotides. Moreover, a combination of single-molecule fluorescence imaging and in situ single-molecule DNA sequencing was used as an "end-point" measurement to detect various histone modifications and genomic positions for individual nucleosomes (39, 40). Importantly, this technically complex approach did not inform on single-molecule dynamics, and commercial solutions for such in situ single-molecule sequencing are not yet readily available. To our knowledge, multiplex measurements where real-time dynamic trajectories can be recorded at the single-molecule level and matched to a large sequence space have never been reported.

Results

To comprehensively profile complex single-molecule dynamics at the library scale, we integrated high-throughput single-molecule fluorescence resonance energy transfer (smFRET) microscopy (41–44) with Illumina next-generation sequencing (NGS) (45, 46) (Fig. 1). To immobilize a library of FRET constructs on the surface of an Illumina MiSeq flow cell (Fig. 1A), we used a 5′ single-stranded DNA overhang and a ligation adapter complementary to the P5 oligonucleotides on the flow cell (Fig. 1B and figs. S1A and S2D). A P7 sequence on the other end enabled bridge amplification upon ligation to a surface P5 oligonucleotide. An Illumina "Read 1" sequence before the variable part of the constructs provided a starting point for the sequencing reaction. To mount the Illumina flow cell onto our motorized objective–based total internal reflection fluorescence (TIRF) microscope, we custom-designed and 3D printed an adapter (Fig. 1C). Using the adapter, we collected smFRET data from a grid of fields of view (FOVs) that together cover the available area of the flow cell. A Peltier element mounted on top of the flow cell allowed us to control the temperature. Next, we subjected the flow cell to the standard Illumina sequencing workflow. Bridge amplification was localized to the immediate vicinity (~1 μm) of the template, and individual molecules observed during smFRET imaging could therefore be matched with the corresponding sequenced clusters according to their positions on the flow cell surface (fig. S3). The final output of our multiplexed single-molecule characterization at the library scale (MUSCLE) workflow consists of smFRET time trajectories with handles that identify the underlying DNA sequences.

MUSCLE reveals sequence-dependent DNA hairpin properties

To demonstrate the capabilities of MUSCLE, we used it to study DNA hairpins—single-stranded DNA structures formed by two complementary regions separated by a noncomplementary loop. DNA hairpins have been implicated in many biological processes, including replication, transcription, and recombination (47, 48), and are also of central importance to biotechnology (49–54). Hairpins serve as prime model systems to study nucleic acid hybridization, one of the most fundamental processes that control the flow of genetic information in the cell (55, 56). Hairpins typically show two-state folding kinetics, which has been visualized for a small number of sequences by smFRET (57–67).

To comprehensively profile the sequence dependence of DNA hairpin properties, our FRET constructs featured Cy3 donor and Cy5 acceptor fluorophores separated by a DNA hairpin with an 18-nucleotide (nt) polyA loop and a 6 to 8–base pair (bp) stem across a 256-member library (Fig. 1D). We note that the placement and type of fluorophores can affect DNA hairpin properties (65, 68). In our design, the 5′ and 3′ terminal fluorophores were separated from the hairpin stem by three unpaired nucleotides on each side and were located far from the variable region. The fluorophores were therefore unlikely to exert substantial effects on stem hybridization. However, even if such an effect were present, it would be extremely unlikely to be sequence specific and would therefore not affect relative comparisons between different library members. With this labeling scheme, the closed and open conformations of the hairpin resulted in high (~0.8) and low (~0.2) FRET efficiencies, respectively (fig. S4). Time trajectories from individual DNA hairpins displayed spontaneous transitions between the two FRET states, with the time spent in the closed state increasing with stem length (fig. S4). We collected MUSCLE datasets at 10°, 15°, and 20°C (Figs. 1E and 2, A to D, and figs. S1 and S5).

The equilibrium constants and Gibbs free energies for hairpin formation provided a comprehensive overview of hairpin stability as a function of sequence and temperature (Fig. 2,

[1]Department of Cell and Molecular Biology, Science for Life Laboratory, Uppsala University, 75105 Uppsala, Sweden. [2]Department of Medical Sciences, Science for Life Laboratory, Uppsala University, 75144 Uppsala, Sweden. [3]Department of Biochemistry, University of Zürich, 8057 Zürich, Switzerland. [4]Small Biosystems Lab, Condensed Matter Physics Department, Universitat de Barcelona, 08028 Barcelona, Spain. [5]Institut de Nanociència i Nanotecnologia (IN2UB), Universitat de Barcelona, 08028 Barcelona, Spain.
*Corresponding author. Email: sebastian.deindl@icm.uu.se
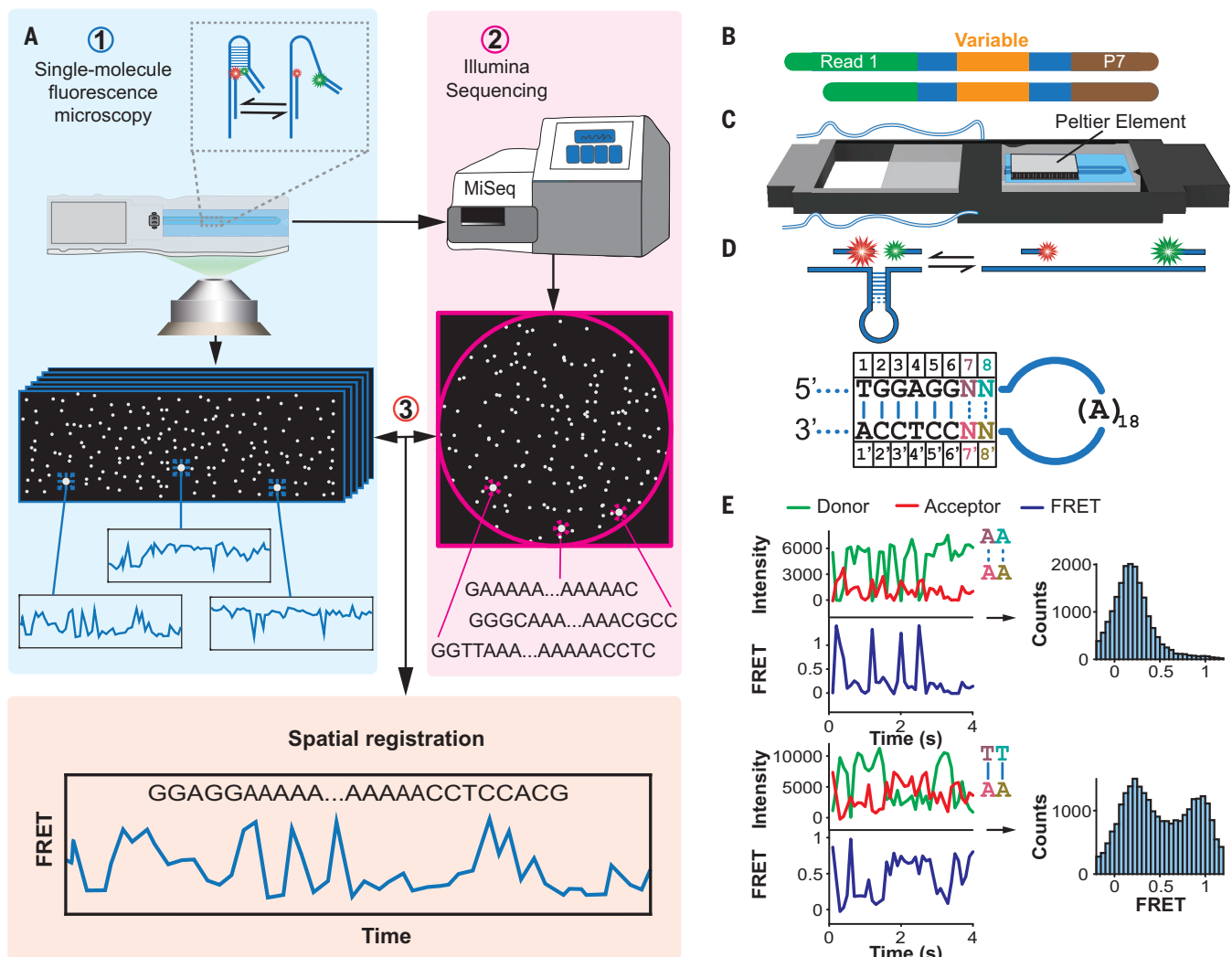†These authors contributed equally to this work.

Fig. 1. General outline of the MUSCLE approach. (**A**) Schematic of the experimental procedure using a DNA hairpin construct as an example. (**B**) General schematic of a MUSCLE library. (**C**) Schematic for smFRET imaging on a temperature-controlled Illumina MiSeq flow cell using a custom-designed microscopy adapter. (**D**) Design of the hairpin construct library with four variable nucleotides. See figs. S1A and S2B and table S1 for more details. (**E**) Example intensity (Cy3, green; Cy5, red) and FRET (blue) time trajectories (left) and FRET efficiency histograms (right) for two individual members of the library that form a 6-bp (top) or 8-bp stem (bottom).

B and C). At all temperatures, the stability of the hairpin was higher along the diagonal of the heatmap, corresponding to Watson-Crick base pairing between the variable nucleotides. With increasing temperatures, the FRET time traces were overall more dynamic with a more populated low-FRET state. Moreover, hairpin stability decreased with increasing temperatures throughout the measured sequence space. As expected, the number of complementary base pairs in the stem was the main contributor to the sequence-dependent stability of the hairpin, with a single match in position 7 favored over one in position 8 when preceded by a mismatch in position 7 (Fig. 2D). The Gibbs free energies obtained from our MUSCLE data displayed an overall good correlation (0.64) with the stabilities as predicted by the nearest-neighbor model mFold (Fig. 2E) (*69, 70*). We

note that to optimally observe differences in stability between library members at moderate temperatures, we carried out our MUSCLE experiments at pH 9, which mFold cannot account for. The nearest-neighbor model appeared to substantially underestimate the stability of hairpins that can form noncanonical base pairs (most notably G-T) and/or contain single-nucleotide mismatches (Fig. 2E, red dots).

Cas9-induced target DNA unwinding-rewinding dynamics are sensitive to PAM-distal complementarity

To further demonstrate the general utility of our method, we used MUSCLE to probe Cas9-induced target DNA unwinding-rewinding dynamics. Cas9, a programmable nuclease, has proven invaluable in research and biotechnology, with particular significance in gene

editing applications (*71, 72*). Nonetheless, despite its inherent precision in targeted DNA cleavage, Cas9 may still induce cleavage at genomic DNA sites with imperfect complementarity to the guide RNA, leading to off-target editing (*73–77*). When considering the therapeutic use of Cas9 enzymes, such off-target activity raises safety concerns. Off-target DNA binding and cleavage are affected by the type, positioning, and total number of mismatches (*30, 73, 75, 78–81*). Although protospacer adjacent motif (PAM)–distal mismatches are compatible with stable binding, they often result in the formation of a catalytically incompetent complex (*79–85*). A reversible unwinding of the target DNA duplex in the PAM-distal region triggers a conformational change in the nuclease and serves as one of the main checkpoints for cleavage (*86*). This process is sensitive to the
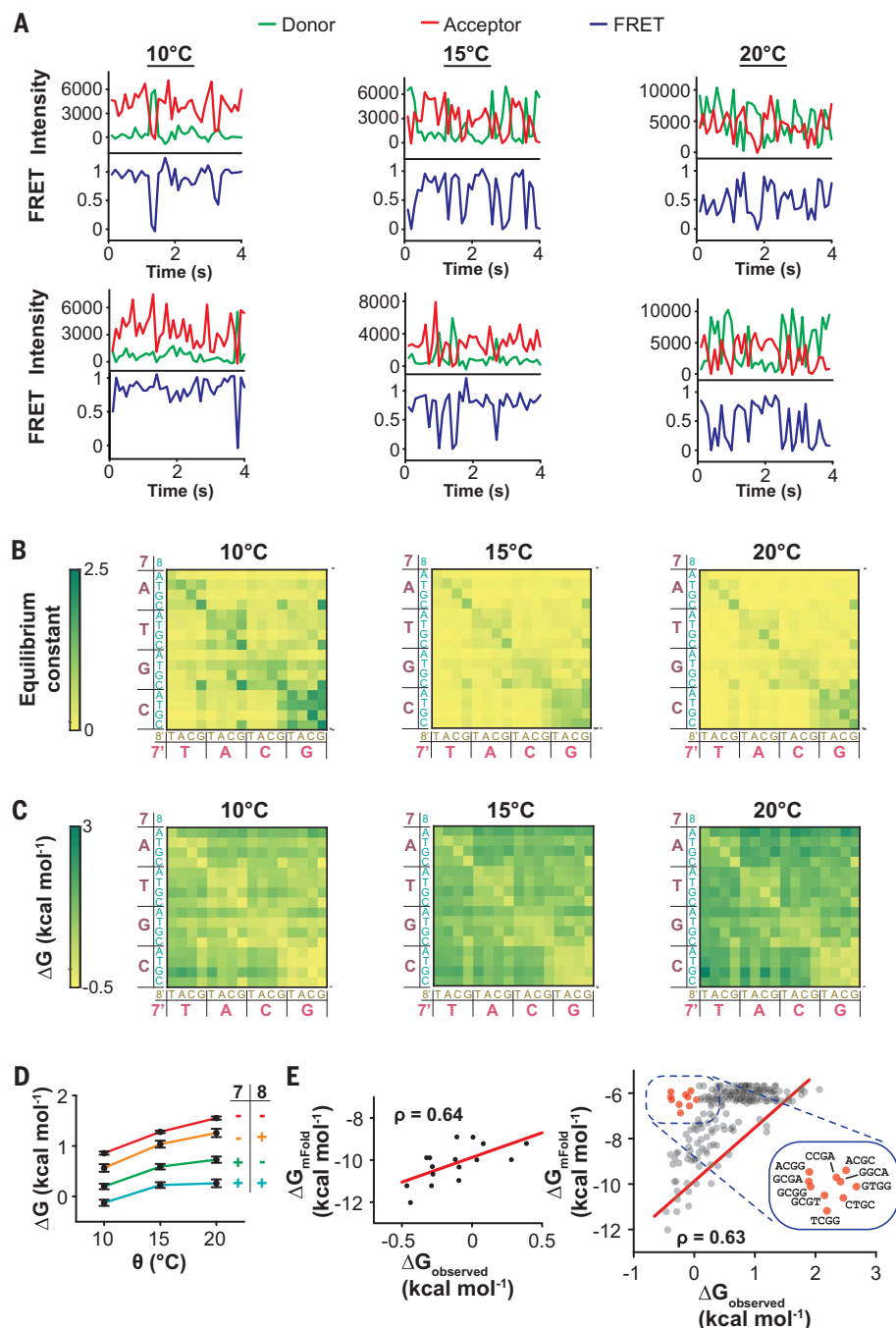
Fig. 2. MUSCLE measurement of hybridization dynamics for a library of DNA hairpins with variable stem sequence. (**A**) Example intensity (Cy3, green; Cy5, red) and FRET (blue) time trajectories for a library member TTAA that can form two additional A-T base pairs at 10°C (left), 15°C (center), and 20°C (right). Heatmaps of equilibrium constants (**B**) and Gibbs free energies (**C**) for all library members at 10°C (left), 15°C (center), and 20°C (right). Complementary nucleotides are located on the diagonal. (**D**) Plots of average Gibbs free energies versus temperature for all library members with additional complementary base pairs in none (red), only second (orange), only first (green), or both (blue) variable positions. Error bars represent standard errors of the mean. (**E**) Comparison of measured Gibbs free energies with mFold predictions at 10°C for all library members (right) or only those that form two additional complementary base pairs (left). In each case, ρ denotes the Pearson correlation coefficient. The red lines in both plots represent the linear fit of the data for library members that form two additional complementary base pairs. Sequences that are substantially more stable than expected on the basis of the nearest-neighbor prediction are highlighted in red. N = 124,387, 139,838, and 99,538 molecules for 10°, 15°, and 20°C, respectively.

sequence complementarity in the PAM-distal region and plays a key role in the improved specificity of Cas9 variants. At the same time, structural variations in target DNA-RNA interactions can have a strong effect on the unwinding-rewinding dynamics (*75*). However, the lack of suitable methodology has thus far prevented an exhaustive analysis of these dynamics across the sequence space. We therefore adapted a smFRET assay (*86*) to monitor Cas9-induced target DNA unwinding-rewinding dynamics using MUSCLE (Fig. 3 and fig. S6A).

The library constructs featured donor and acceptor dyes on opposite strands of the target DNA and spanned all combinations of nucleotides in positions 17 to 20, while preserving complementarity of the DNA duplex (Fig. 3A and fig. S2C). Upon addition of catalytically inactive Cas9 (dCas9) at saturating concentrations (fig. S8), the initially-high FRET (fig. S9A) decreased (Fig. 3, B and C, and fig. S9B) owing to the formation of the DNA-RNA hybrid in the PAM-distal region. For the perfectly matching target DNA, time traces displayed continuously low FRET values, indicating stable unwinding. Conversely, in the presence of mismatches in the target DNA, frequent reversions to higher FRET were observed, indicating a dynamic equilibrium between unwinding and rewinding (Fig. 3C and fig. S7). The time spent in the unwound state, and therefore the equilibrium constant, varied substantially depending on the target sequence (Fig. 3, B and C). In stark contrast, the measurements for each sequence, determined in two independent experiments, displayed a correlation of 0.9, underscoring the robust repeatability of MUSCLE measurements (fig. S6D).

Notably, the number and positions of mismatches were among the main factors affecting unwinding (Fig. 3, D and E). Mismatches where the target DNA base was rendered identical to the corresponding RNA base consistently had the most disruptive effect (fig. S6B), likely because of the inability to form noncanonical base pairs (*73, 75, 78*). We found that some sequences exhibited a notably shifted low-FRET peak, potentially indicating variability in the conformation of the unwound state (fig. S9, B to D).

To estimate the accuracy of matching sequenced clusters to respective smFRET time trajectories, we added a small fraction (<1%) of a control construct with a distinct sequence and starting FRET value to a MUSCLE experiment (fig. S6C). Analysis of starting FRET distributions for the control and library constructs sorted on the basis of sequence demonstrated that at least 96% of smFRET time trajectories were matched with the correct sequence. As the surface of Illumina flow cells is not optimized for single-molecule fluorescence imaging, we compared smFRET time trajectories for fully matched DNA targets

obtained using MUSCLE with those from conventional smFRET under similar conditions. Despite a higher background in the Cy5 channel upon excitation with a 532-nm laser, the signal-to-noise ratio for MUSCLE traces was comparable to those obtained by conventional smFRET microscopy on a PEGylated quartz slide (fig. S6E). High-quality smFRET time trajectories can thus be readily obtained on the surface of an Illumina flow cell.

The single-molecule capability of our method enabled us to derive kinetic rates from direct equilibrium observations. Consequently, we used hidden Markov model (HMM) analysis on the FRET time traces to determine the sequence-dependent rates for Cas9-mediated unwinding of target DNA (Fig. 3, F and G, and fig. S10). As expected, the complementarity between the guide RNA and the target DNA primarily affected the rewinding rate, whereas the G-C content of the target exerted a stronger influence on the unwinding rate (fig. S10, C and D). Notably, the rewinding rate exhibited a substantially stronger correlation with the mean FRET than the unwinding rate (0.93 versus 0.7), indicating that the rewinding rate is the main factor that determines the equilibrium between target DNA unwinding and rewinding (Fig. 3G). This observation is consistent with the fact that the variation in stability among fully complementary target DNA duplexes was considerably lower compared with that among DNA-RNA hybrids, which can have between zero and four mismatches.

DNA unwinding-rewinding dynamics correlate with Cas9 cleavage rates

To examine the relationship between the target DNA unwinding and Cas9-mediated cleavage activity, we turned to a gel-based cleavage assay and tested 16 target DNA sequences that span the range of observed equilibrium and dynamic behaviors (fig. S12, A and B). We observed a variation of at least 30-fold in cleavage rates across the sequences tested (fig. S12, A and B). To assess cleavage rates for all library members, we leveraged an NGS-based cleavage assay (Fig. 4A). Results from both assays displayed strong correlation (0.78), with the dynamic range of the gel-based assay substantially exceeding that of the NGS-based assay (fig. S12B). Overall, the DNA sequences that spend more time in the unwound state displayed more efficient cleavage (Fig. 4B, correlation coefficient: 0.66), consistent with an important role of the unwinding-rewinding equilibrium in Cas9 activity (75, 82–89). At the same time, a number of sequences substantially deviated from this general trend (fig. S12D). As all the target sequences had at least 16 base pairs of complementarity with the guide RNA, variations in binding affinity are unlikely to underlie these deviations. MUSCLE analysis therefore could pinpoint sequences
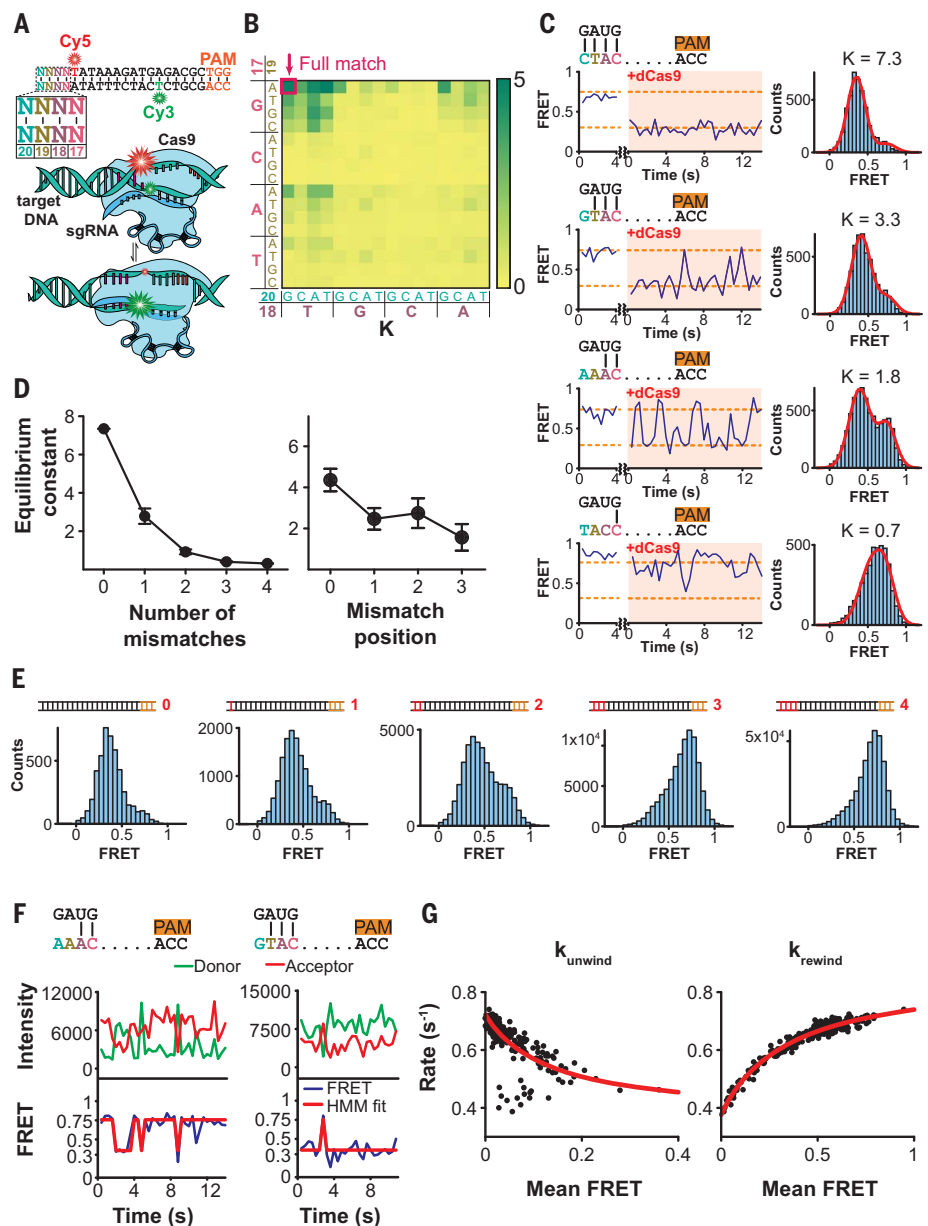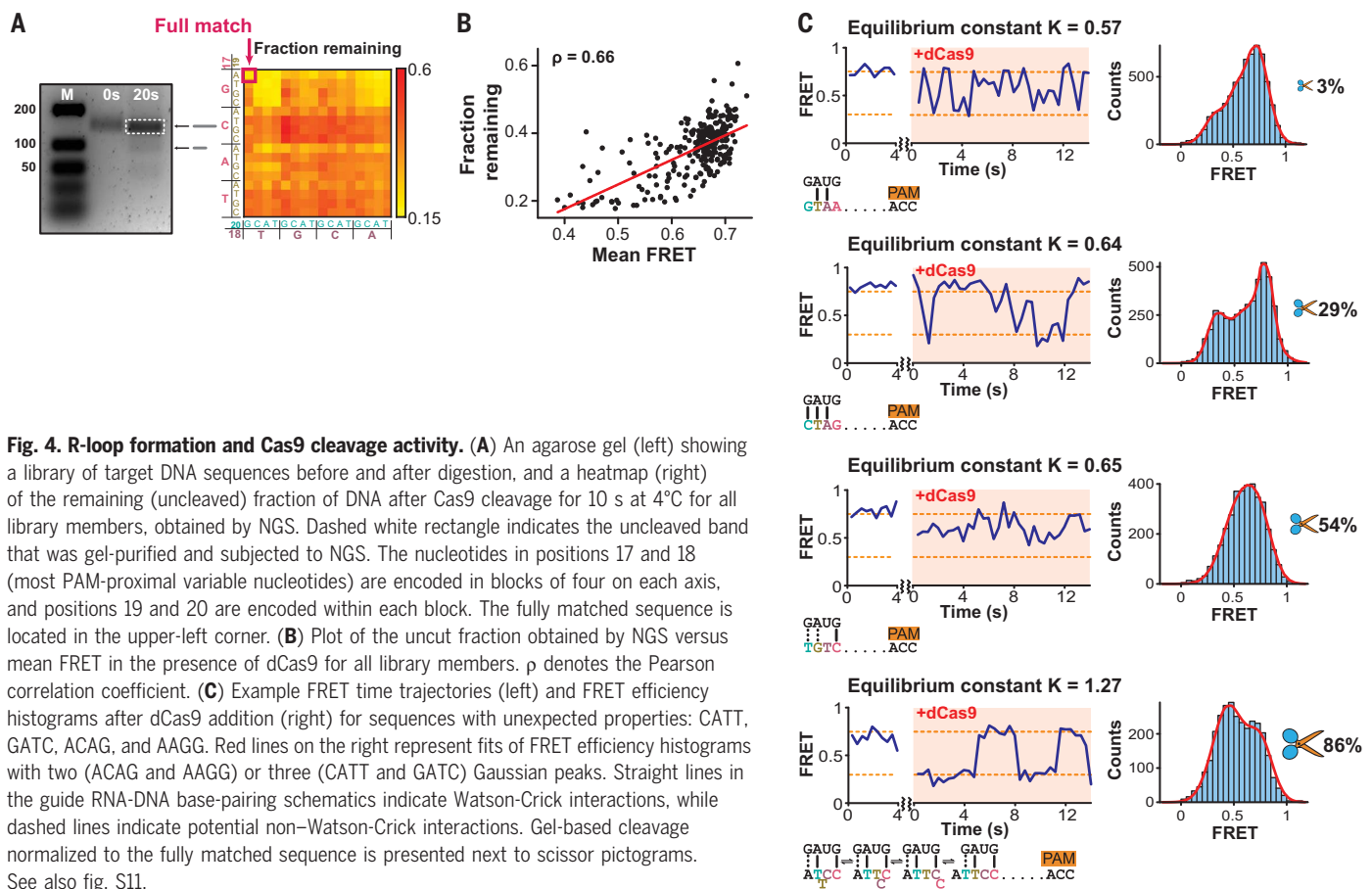

Fig. 3. MUSCLE measurements of Cas9-induced target DNA unwinding-rewinding dynamics for a library with variable PAM-distal sequence. (**A**) Schematic of the FRET assay and design of the construct library with four variable nucleotides in the PAM-distal region. See fig. S2C and table S1 for more details. (**B**) Heatmap of the equilibrium constant for all library members in the presence of dCas9. The nucleotides in positions 17 and 18 (most PAM-proximal variable nucleotides) are encoded in blocks of four on each axis, and positions 19 and 20 are encoded within each block. The fully matched sequence is located in the upper-left corner. $N = 125,697$ molecules. (**C**) Example FRET time trajectories (left) and FRET efficiency histograms after dCas9 addition (right) for four individual members of the library that have between zero and three mismatches. See also fig. S7. (**D**) Plot of the equilibrium constant as a function of the number of consecutive mismatches starting from the PAM-distal target DNA end (left) or of the position of a single mismatch (right). Error bars represent standard errors of the mean. (**E**) FRET efficiency histograms after dCas9 addition for library members combined based on the number of consecutive mismatches (red) starting from the PAM-distal (orange) target DNA end. (**F**) Example intensity (Cy3, green; Cy5, red) and FRET (blue) time trajectories, and Viterbi paths (red) obtained using HMM analysis for two library members. (**G**) Plots of unwinding (left) and rewinding (right) rates versus mean FRET values for individual library members (black dots) and fit results with a model assuming a constant rewinding or unwinding rate, respectively (red lines).

Fig. 4. R-loop formation and Cas9 cleavage activity. (**A**) An agarose gel (left) showing a library of target DNA sequences before and after digestion, and a heatmap (right) of the remaining (uncleaved) fraction of DNA after Cas9 cleavage for 10 s at 4°C for all library members, obtained by NGS. Dashed white rectangle indicates the uncleaved band that was gel-purified and subjected to NGS. The nucleotides in positions 17 and 18 (most PAM-proximal variable nucleotides) are encoded in blocks of four on each axis, and positions 19 and 20 are encoded within each block. The fully matched sequence is located in the upper-left corner. (**B**) Plot of the uncut fraction obtained by NGS versus mean FRET in the presence of dCas9 for all library members. ρ denotes the Pearson correlation coefficient. (**C**) Example FRET time trajectories (left) and FRET efficiency histograms after dCas9 addition (right) for sequences with unexpected properties: CATT, GATC, ACAG, and AAGG. Red lines on the right represent fits of FRET efficiency histograms with two (ACAG and AAGG) or three (CATT and GATC) Gaussian peaks. Straight lines in the guide RNA-DNA base-pairing schematics indicate Watson-Crick interactions, while dashed lines indicate potential non–Watson-Crick interactions. Gel-based cleavage normalized to the fully matched sequence is presented next to scissor pictograms. See also fig. S11.

with the rate-limiting step for Cas9 cleavage differing from Cas9 binding or DNA unwinding.

Several target sequences exhibited unexpected behaviors (Fig. 4C and fig. S11). Targets with the nucleotides GATC or CATT in the variable region (versus GATG for the fully matched target) were cleaved much more slowly than expected (29 or 3% cleavage in the gel-based assay, respectively, compared with the full match), despite having three or two matches, respectively. Even though the GATC target had a mismatch in position 17, which is the most critical for cleavage out of the four variable positions (79, 80, 82, 84, 86), cleavage was still substantially reduced in comparison to GATA with a mismatch in the same position (29% versus 53% cleavage compared with the full match). Because the folding of the nontarget strand has been implicated in Cas9 activity (90–93), we analyzed the potential to form secondary structures for these sequences. We found that both sequences could form unusually stable loops in the nontarget strand (fig. S12C). These structures could interfere with the unwinding-rewinding equilibrium or hinder cleavage, and they explain the apparent discrepancy between the relatively efficient target DNA unwinding for GATC and CATT and their slow cleavage. Indeed, FRET time trajectories and distributions sug-

gested the existence of an intermediate FRET state for these sequences (Fig. 4C), consistent with a distorted conformation that accommodates a loop in the nontarget strand and hinders cleavage.

The target with ACAG in positions 17 to 20 demonstrated unexpectedly efficient cleavage (54% compared with the full match), despite only featuring a single match in position 17 (Fig. 4C). The elevated activity could be due to noncanonical base pairing interactions, including one rG-dT pair, which is especially well tolerated by Cas9 (80, 94–96). We found that the AAGG sequence with two mismatches displayed a cleavage rate and mean FRET value close to those of the fully matched target (GATG). Although noncanonical base pairing might partially contribute to this, we noticed that this sequence could form at least three alternative duplex structures, each featuring two canonical base pairs and a bulge in the target strand (Fig. 4C). We speculate that this sequence therefore entropically favors the unwound conformation, which shifts the unwinding-rewinding equilibrium and thereby increases the cleavage rate.

Target sequence can substantially affect Cas9 specificity (73, 74, 97, 98). To assess whether the observed effects of mismatches in positions 17 to 20 were specific to the target sequence, we

repeated the NGS-based cleavage assay for a different library where nucleotides 1 to 16 in the PAM-proximal region were replaced with a segment from the GEMIN5 gene (79, 99), while nucleotides 17 to 20 were variable, as before. The single-guide RNA for this library was designed to target the GEMIN5-derived sequence at positions 1 to 16, while maintaining the same sequence (GATG) as the original library at variable positions 17 to 20 to facilitate the comparison. The cleavage values for the two libraries (fig. S13) exhibited a strong correlation (0.82), indicating that the observed patterns of cleavage efficiency are more general and not specific to a particular target sequence context.

Conclusions

We have developed a method, MUSCLE, that combines single-molecule fluorescence microscopy with next-generation sequencing to enable highly multiplexed observations of complex dynamics. Our data provide mechanistic insights into the off-target activity of Cas9 and can be used to refine guide efficacy and off-target prediction algorithms. We envision that MUSCLE will contribute to the development of improved RNA-guided nucleases. A primary benefit of a comprehensive analysis lies in its independence from a priori hypotheses, rendering it considerably more effective in the detection of

unforeseen phenomena. Indeed, MUSCLE allowed us to discover and scrutinize "outlier" target sequences with unexpected dynamic behaviors. Given that both fluorescence microscopy and MiSeq instruments are ubiquitous and the adapter can be simply 3D printed, we anticipate that MUSCLE will be readily adopted by the scientific community. Our approach lends itself to the study of not only a wide range of proteins acting on libraries of modified or unmodified nucleic acids but also DNA-barcoded proteins, compounds, or ligands. We note that longer library constructs can also be used in the MUSCLE approach by designing them to be cleaved into shorter (<300 bp) DNA molecules for amplification and sequencing at the end of a single-molecule experiment. In this scenario, the variable sequence portion does not have to be confined to the remaining construct after cleavage, as barcoding can encode sequence variability elsewhere. Moreover, we envision that MUSCLE can be implemented with other single-molecule techniques, such as force spectroscopy. For example, magnetic tweezers offer the possibility to conduct parallel measurements on a large number of molecules (*100–102*). We note that the library size in our approach is limited only by the measurement throughput. For example, a single MUSCLE experiment with a full MiSeq flow cell was sufficient to obtain complete coverage of a 4096-member hairpin library containing six variable nucleotides (fig. S14). We therefore anticipate that much larger libraries (>10,000 members) can be analyzed using MUSCLE, thus paving the way for genome-scale single-molecule biophysics.

REFERENCES AND NOTES

1. S. Mohapatra, C.-T. Lin, X. A. Feng, A. Basu, T. Ha, *Chem. Rev.* **120**, 36–78 (2020).
2. L. Bacic, A. Sabantsev, S. Deindl, *Curr. Opin. Struct. Biol.* **65**, 61–68 (2020).
3. F. R. Hill, E. Monachino, A. M. van Oijen, *Biochem. Soc. Trans.* **45**, 759–769 (2017).
4. J. Choi, R. Grosely, E. V. Puglisi, J. D. Puglisi, *Curr. Opin. Struct. Biol.* **58**, 233–240 (2019).
5. M. Orrit, T. Ha, V. Sandoghdar, *Chem. Soc. Rev.* **43**, 973–976 (2014).
6. F. Kulzer, M. Orrit, *Annu. Rev. Phys. Chem.* **55**, 585–611 (2004).
7. M. Dangkulwanich, T. Ishibashi, L. Bintu, C. Bustamante, *Chem. Rev.* **114**, 3203–3223 (2014).
8. J. Zhou, V. Schweikhard, S. M. Block, *Biochim. Biophys. Acta Gene Regul. Mech.* **1829**, 29–38 (2013).
9. J. Michaelis, B. Treutlein, *Chem. Rev.* **113**, 8377–8399 (2013).
10. L. Bai, T. J. Santangelo, M. D. Wang, *Annu. Rev. Biophys. Biomol. Struct.* **35**, 343–360 (2006).
11. M. Stracy, A. N. Kapanidis, *Methods* **120**, 103–114 (2017).
12. E. Lerner *et al.*, *Science* **359**, eaan1133 (2018).
13. D. Dulin, B. A. Berghuis, M. Depken, N. H. Dekker, *Curr. Opin. Struct. Biol.* **34**, 116–122 (2015).
14. M. F. Juette *et al.*, *Curr. Opin. Chem. Biol.* **20**, 103–111 (2014).
15. J. H. Felce, S. J. Davis, D. Klenerman, *Trends Pharmacol. Sci.* **39**, 96–108 (2018).
16. J. Camunas-Soler, M. Ribezzi-Crivellari, F. Ritort, *Annu. Rev. Biophys.* **45**, 65–84 (2016).
17. S. Park, G. B. Brandani, T. Ha, G. D. Bowman, *Nucleic Acids Res.* **51**, 10326–10343 (2023).
18. A. A. Hoskins, J. Gelles, M. J. Moore, *Curr. Opin. Chem. Biol.* **15**, 864–870 (2011).
19. J. M. Tome *et al.*, *Nat. Methods* **11**, 683–688 (2014).
20. R. Nutiu *et al.*, *Nat. Biotechnol.* **29**, 659–664 (2011).
21. C. Jung *et al.*, *Cell* **170**, 35–47.e13 (2017).
22. R. E. Kleiner, C. E. Dumelin, G. C. Tiu, K. Sakurai, D. R. Liu, *J. Am. Chem. Soc.* **132**, 11779–11791 (2010).
23. A. Basu *et al.*, *Nature* **589**, 462–467 (2021).
24. U. T. T. Nguyen *et al.*, *Nat. Methods* **11**, 834–840 (2014).
25. G. P. Dann *et al.*, *Nature* **548**, 607–611 (2017).
26. I. Severins, C. Joo, J. van Noort, *Mol. Cell* **82**, 1788–1805 (2022).
27. J. B. Kinney, D. M. McCandlish, *Annu. Rev. Genomics Hum. Genet.* **20**, 99–127 (2019).
28. E. Marklund, Y. Ke, W. J. Greenleaf, *Nat. Rev. Genet.* **24**, 401–414 (2023).
29. C. J. Layton, P. L. McMahon, W. J. Greenleaf, *Mol. Cell* **73**, 1075–1082.e4 (2019).
30. E. A. Boyle *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **114**, 5461–5466 (2017).
31. N. Svensen, O. B. Peersen, S. R. Jaffrey, *ChemBioChem* **17**, 1628–1635 (2016).
32. L. C. Szymczak, H.-Y. Kuo, M. Mrksich, *Anal. Chem.* **90**, 266–282 (2018).
33. D. Wu *et al.*, *Anal. Chem.* **95**, 2645–2652 (2023).
34. Z. Li *et al.*, *Sci. Adv.* **6**, eabb3350 (2020).
35. N. Mamet *et al.*, bioRxiv 630830 [Preprint] (2019); https://doi.org/10.1101/630830.
36. J. D. Buenrostro *et al.*, *Nat. Biotechnol.* **32**, 562–568 (2014).
37. R. Andrews *et al.*, bioRxiv 2022.02.27.482175 [Preprint] (2022); https://doi.org/10.1101/2022.02.27.482175.
38. K. Makasheva *et al.*, *J. Am. Chem. Soc.* **143**, 16313–16319 (2021).
39. E. Shema *et al.*, *Science* **352**, 717–721 (2016).
40. V. Fedyuk *et al.*, *Nat. Biotechnol.* **41**, 212–221 (2023).
41. T. Ha *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 6264–6268 (1996).
42. T. Ha *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 9077–9082 (1999).
43. X. Zhuang *et al.*, *Science* **288**, 2048–2051 (2000).
44. T. Förster, *Ann. Phys.* **437**, 55–75 (1948).
45. D. R. Bentley *et al.*, *Nature* **456**, 53–59 (2008).
46. S. J. Maerkl, S. R. Quake, *Science* **315**, 233–237 (2007).
47. D. Bikard, C. Loot, Z. Baharoglu, D. Mazel, *Microbiol. Mol. Biol. Rev.* **74**, 570–588 (2010).
48. G. Wang, K. M. Vasquez, *Nat. Rev. Genet.* **24**, 211–234 (2023).
49. T. Omabegho, R. Sha, N. C. Seeman, *Science* **324**, 67–71 (2009).
50. T. E. Tomov *et al.*, *J. Am. Chem. Soc.* **135**, 11935–11941 (2013).
51. Y. Ning, J. Hu, F. Lu, *Biomed. Pharmacother.* **132**, 110902 (2020).
52. T.-H. Ku *et al.*, *Sensors* **15**, 16281–16313 (2015).
53. S. J. Green, D. Lubrich, A. J. Turberfield, *Biophys. J.* **91**, 2966–2975 (2006).
54. V. V. Sinitsyna, A. A. Vetcher, *Biomedicines* **10**, 1079 (2022).
55. L. E. Baltierra-Jasso, M. J. Morten, L. Laflör, S. D. Quinn, S. W. Magennis, *J. Am. Chem. Soc.* **137**, 16020–16023 (2015).
56. J. M. Huguet, M. Ribezzi-Crivellari, C. V. Bizarro, F. Ritort, *Nucleic Acids Res.* **45**, 12921–12931 (2017).
57. D. Bandyopadhyay, P. P. Mishra, *Front. Mol. Biosci.* **8**, 725541 (2021).
58. S. Patra *et al.*, *Nucleic Acids Res.* **47**, 981–996 (2019).
59. J. R. Grunwell *et al.*, *J. Am. Chem. Soc.* **123**, 4295–4303 (2001).
60. E. Nir *et al.*, *J. Phys. Chem. B* **110**, 22103–22124 (2006).
61. T. E. Tomov *et al.*, *Biophys. J.* **102**, 1163–1173 (2012).
62. Y. Santoso, J. P. Torella, A. N. Kapanidis, *ChemPhysChem* **11**, 2209–2219 (2010).
63. J. P. Torella, S. J. Holden, Y. Santoso, J. Hohlbein, A. N. Kapanidis, *Biophys. J.* **100**, 1568–1577 (2011).
64. M. I. Wallace, L. Ying, S. Balasubramanian, D. Klenerman, *J. Phys. Chem. B* **104**, 11551–11555 (2000).
65. W. Kügel, A. Muschielok, J. Michaelis, *ChemPhysChem* **13**, 1013–1022 (2012).
66. M. L. Mitchell, M. P. Leveille, R. S. Solecki, T. Tran, B. Cannon, *J. Phys. Chem. B* **122**, 11841–11851 (2018).
67. R. Tsukanov *et al.*, *J. Phys. Chem. B* **117**, 11932–11942 (2013).
68. A. Hartmann, G. Krainer, M. Schlierf, *Molecules* **19**, 13735–13754 (2014).
69. M. Zuker, *Nucleic Acids Res.* **31**, 3406–3415 (2003).
70. J. SantaLucia Jr., *Proc. Natl. Acad. Sci. U.S.A.* **95**, 1460–1465 (1998).
71. M. Jinek *et al.*, *Science* **337**, 816–821 (2012).
72. A. V. Anzalone, L. W. Koblan, D. R. Liu, *Nat. Biotechnol.* **38**, 824–844 (2020).
73. P. D. Hsu *et al.*, *Nat. Biotechnol.* **31**, 827–832 (2013).
74. V. Pattanayak *et al.*, *Nat. Biotechnol.* **31**, 839–843 (2013).
75. M. Pacesa *et al.*, *Cell* **185**, 4067–4081.e21 (2022).
76. S. Q. Tsai *et al.*, *Nat. Biotechnol.* **33**, 187–197 (2015).
77. P. Cameron *et al.*, *Nat. Methods* **14**, 600–606 (2017).
78. J. G. Doench *et al.*, *Nat. Biotechnol.* **34**, 184–191 (2016).
79. E. A. Boyle *et al.*, *Sci. Adv.* **7**, eabe5496 (2021).
80. S. K. Jones Jr. *et al.*, *Nat. Biotechnol.* **39**, 84–93 (2021).
81. L. Zhang *et al.*, *Nucleic Acids Res.* **48**, 5037–5053 (2020).
82. Y. S. Dagdas, J. S. Chen, S. H. Sternberg, J. A. Doudna, A. Yildiz, *Sci. Adv.* **3**, eaao0027 (2017).
83. I. E. Ivanov *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **117**, 5853–5860 (2020).
84. S. H. Sternberg, B. LaFrance, M. Kaplan, J. A. Doudna, *Nature* **527**, 110–113 (2015).
85. M. Yang *et al.*, *Cell Rep.* **22**, 372–382 (2018).
86. D. Singh *et al.*, *Nat. Struct. Mol. Biol.* **25**, 347–354 (2018).
87. M. Pacesa *et al.*, *Nature* **609**, 191–196 (2022).
88. S. Gong, H. H. Yu, K. A. Johnson, D. W. Taylor, *Cell Rep.* **22**, 359–371 (2018).
89. M. D. Newton *et al.*, *Nat. Struct. Mol. Biol.* **26**, 185–192 (2019).
90. F. Jiang *et al.*, *Science* **351**, 867–871 (2016).
91. A. Santiago-Frangos, T. Wiegand, B. Wiedenheft, *EMBO J.* **38**, e101474 (2019).
92. H. Balci, V. Globyte, C. Joo, *ACS Chem. Biol.* **16**, 596–603 (2021).
93. G. Palermo, Y. Miao, R. C. Walker, M. Jinek, J. A. McCammon, *ACS Cent. Sci.* **2**, 756–763 (2016).
94. N. Sugimoto, I. Yasumatsu, M. Fujimoto, *Nucleic Acids Symp. Ser.* **37**, 199–200 (1997).
95. N. Sugimoto, M. Nakano, S. Nakano, *Biochemistry* **39**, 11270–11281 (2000).
96. I. J. Kimsey, K. Petzold, B. Sathyamoorthy, Z. W. Stein, H. M. Al-Hashimi, *Nature* **519**, 315–320 (2015).
97. X. Wu, A. J. Kriz, P. A. Sharp, *Quant. Biol.* **2**, 59–70 (2014).
98. Y. Fu *et al.*, *Nat. Biotechnol.* **31**, 822–826 (2013).
99. H. Xu *et al.*, *Genome Res.* **25**, 1147–1157 (2015).
100. I. De Vlaminck *et al.*, *Nano Lett.* **11**, 5489–5493 (2011).
101. N. Ribeck, O. A. Saleh, *Rev. Sci. Instrum.* **79**, 094301 (2008).
102. D. Dulin *et al.*, *Biophys. J.* **109**, 2113–2125 (2015).
103. J. Aguirre Rivera *et al.*, Massively parallel analysis of single-molecule dynamics on next generation sequencing chips, SciLifeLab Data Repository (2024); https://doi.org/10.17044/scilifelab.25705512.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adn5371
Materials and Methods
Figs. S1 to S16
Table S1
References (*104–116*)
MDAR Reproducibility Checklist
Data S1