Helicase Stepping Investigated with One-Nucleotide Resolution Fluorescence Resonance Energy Transfer

Wenxia Lin,[1,4] Jianbing Ma,[1,4] Daguan Nong,[1,4] Chunhua Xu,[1,4] Bo Zhang,[2] Jinghua Li,[1] Qi Jia,[1,4] Shuoxing Dou,[1,4] Fangfu Ye,[1,4] Xuguang Xi,[2,3] Ying Lu,[1,4,*] and Ming Li[1,4,†]

[1]*Beijing National Laboratory for Condensed Matter Physics and CAS Key Laboratory of Soft Matter Physics,
Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China*
[2]*College of Life Sciences, Northwest A&F University, Yangling, Shaanxi 712100, China*
[3]*LBPA, ENS de Cachan, CNRS, Université Paris-Saclay, F-94235 Cachan, France*
[4]*School of Physical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China*

Single-molecule Förster resonance energy transfer is widely applied to study helicases by detecting distance changes between a pair of dyes anchored to overhangs of a forked DNA. However, it has been lacking single-base pair (1-bp) resolution required for revealing stepping kinetics of helicases. We designed a nanotensioner in which a short DNA is bent to exert force on the overhangs, just as in optical or magnetic tweezers. The strategy improved the resolution of Förster resonance energy transfer to 0.5 bp, high enough to uncover differences in DNA unwinding by yeast Pif1 and *E. coli* RecQ whose unwinding behaviors cannot be differentiated by currently practiced methods. We found that Pif1 exhibits 1-bp-stepping kinetics, while RecQ breaks 1 bp at a time but sequesters the nascent nucleotides and releases them randomly. The high-resolution data allowed us to propose a three-parameter model to quantitatively interpret the apparently different unwinding behaviors of the two helicases which belong to two superfamilies.

Helicases are motor proteins involved in almost every aspect of nucleic acid metabolism [1–4]. When a helicase is loaded onto one of the overhangs (i.e., the tracking strand) of a forked DNA duplex and unwinds it, two nucleotides are released per base pair unwound, leading to an increase in end-to-end distance of the overhangs. It is of great interest to know how a helicase uses the discrete energy derived from nucleoside triphosphates (NTPs) hydrolysis to unwind DNA. The question remained unanswered for most helicases due to the lack of methods that can interrogate the stepping kinetics of helicases [5–7]. Optical tweezers (OTs) are so far the most reliable technique to study helicases with 0.5-bp resolution [8–10]. However, OT instrumentation is complicated so that it is accessible to few laboratories only. Magnetic tweezers can measure tens of molecules at a time [11]. Using a bright laser source for illumination and tracking at ∼10 kHz, Dulin *et al.* [12] were able to resolve 0.5 nm steps with a 0.5 s period for double-stranded DNA-tethered beads. The resolution for tethered DNA hairpins is, however, inferior to that for double-stranded DNA of comparable length [12]. Single-molecule Förster resonance energy transfer (smFRET) is a high-throughput technique for helicase assays. Using wide-field fluorescence microscopy, one can routinely record signals in parallel from hundreds of single molecules tethered to a surface [13–16]. To date, the resolution of smFRET is limited to 2–3 bp when forked DNA is used as the substrate. The inter-dye distances are significantly reduced because of thermal fluctuation of the overhangs [17] [Fig. 1(a)]. In addition, the negatively charged ssDNA

may become unstable in the presence of multiple valence cations and is also vulnerable to random disturbances in the fluid chamber caused by, for instance, environment noises. These may result in Förster resonance energy transfer (FRET) traces with large fluctuations (Figs. S1 and S2 in the Supplemental Material [18]). If tension is exerted on the overhangs, the distance between the two dyes can be stabilized and increased. As a result, the resolution of smFRET would be improved. The idea inspired combinations of optical or magnetic tweezers with smFRET [23–27].

Recently, Zocchi *et al.* [28,29] invented a smart device to make use of the stiffness of DNA to control protein conformations. Here, we designed a nanotensioner with which a short DNA duplex is bent to stretch overhangs of a forked DNA [Fig. 1(b)]. Because of the mismatch of their lengths, the short DNA duplex is bent in to an arc, exerting force on the fork. We applied the nanotensioner method to assess two helicases, namely, the yeast Pif1 and the *E. coli* RecQ. Pif1 is a prototypical member of the 5′–3′ directed helicase superfamily 1B that is conserved from yeast to human [30–32]. It plays critical roles in the maintenance of telomeric DNA via catalytic inhibition of telomerase [33–36]. RecQ is a member of superfamily 2 helicases [37,38]. It translocates in the 3′ to 5′ direction [39]. RecQ plays an important role in DNA damage response, chromosomal stability maintenance and has a vital role in maintaining genome homeostasis [40–43]. We did not observe significant differences between the unwinding behaviors of the two helicases with the forked DNA substrates as in conventional FRET assays (Figs. S3 and S4 in the
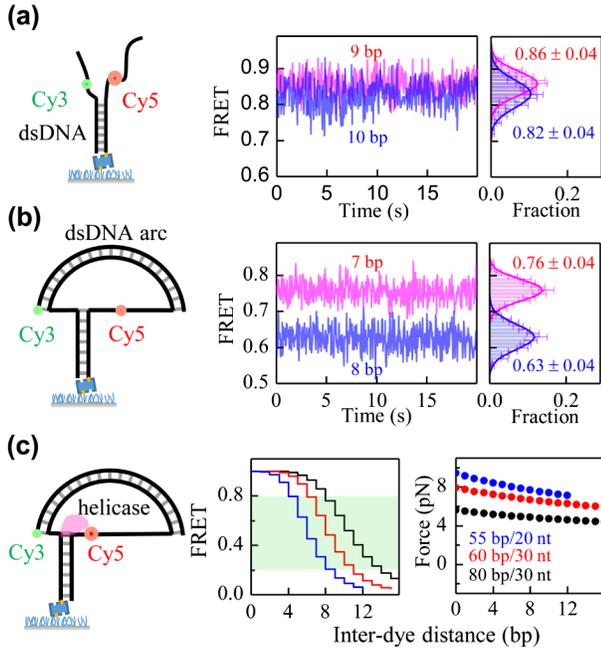
FIG. 1. Design of DNA nanotensioners. (a) Two FRET traces of forked DNA substrates in which Cy3 and Cy5 are 9 bp (pink line) and 10 bp (blue line) distant, respectively. (b) In a nanotensioner, a short dsDNA is bent to exert force on the fork. The FRET efficiency changes significantly when the Cy3-Cy5 distance is increased from 7 bp (pink line) to 8 bp (blue line). The error bars are inversely proportional to the square root of the number of points for each bin. The errors in the Gaussian function fittings are standard deviation (SD). (c) Calculated FRET efficiency (central panel) and tension (right panel) as a function of base pairs unwound by a helicase for various nanotensioners. The green box marks the range of FRET to be analyzed.

Supplemental Material [18]). With the nanotensioner, however, the difference became obvious, enabling us to interrogate the molecular mechanisms of the two helicases with unprecedentedly deep insight.

The design of the nanotensioner [Figs. 1(b) and 1(c)] is based on the fact that dsDNA is semiflexible [44] with a bending modulus of $B \approx 200 \ \mathrm{pN} \times \mathrm{nm}^2$. When a short dsDNA with a contour length S is bent to an arc of radius R, the bending energy is $W = BS/2R^2$. An ssDNA string of length $x = 2R \sin(S/2R)$ maintains the radius of the dsDNA arc by exerting a force F on the two ends of the dsDNA segment. The force can be calculated according to $F = -\delta W/\delta x$, which reads as [29]

$$F = \frac{BS}{R^3}\left[2\sin\left(\frac{S}{2R}\right) - \frac{S}{R}\cos\left(\frac{S}{2R}\right)\right]^{-1}. \quad (1)$$

It is also the tension exerted on the ssDNA overhangs of the forked DNA to be unwound. The harmonic bending model for DNA is valid when the bending energy is lower than the energy of kink formation, 15 to 20 $k_B\mathrm{T}$ as reported [45].

A kink would reduce dramatically the tension on the string. As a result, it is worth noting that, to serve our purpose of studying the unwinding stepping kinetics of a helicase, the DNA construct should meet the following conditions: (i) The tension should not be larger than the one that may create a kink in the arc; (ii) The tension does not decrease too much upon duplex unwinding so that the steps are basically uniform during the experiments. We attached a pair of dyes (Cy3 and Cy5, Förster radius 5.8 nm.) near the junction between the ssDNA overhangs and the duplex of the forked dsDNA [Fig. 1(c)]. Our calculations indicate that the DNA construct in Fig. 1(c) meets the above conditions and works well when the initial length of the string is in the range from 20 to 40 nt and the length of the dsDNA arc ranges from 50 to 100 bp. We show in the central panel of Fig. 1(c) the calculated FRET efficiency versus number of base pairs unwound for three typical nanotensioners. The tension decreases slightly when the string length increases due to the newly added nucleotides [right panel in Fig. 1(c)]. The calculations help us design nanotensioners according to the range of nucleotide distance and the range of force we are interested in. We use a DNA nanotensioner with a 60-bp dsDNA arc and a 30-nt ssDNA string in the following experiments. The force is estimated to be about 6 pN. The FRET change induced by each base pair unwound is about 0.13 [Fig. 1(b)], which is large enough to be readily recorded by many commercial single-molecule fluorescence microscopes. In contrast, the FRET change is only about 0.04 for a forked DNA with free overhangs [Fig. 1(a)].

We first characterized the unwinding kinetics with a forked DNA substrate as in conventional FRET assays. The experiments were performed with objective-based total-internal-reflection fluorescence microscopy (see the Supplemental Material [18] for details). We observed similar unwinding bursts for both helicases in buffers containing high concentration ATP (Fig. S3 in Supplemental Material [18]). We then reduced the ATP concentration to 0.5–2 μM in order to see possible unwinding steps [14,15]. Unfortunately, stepping events for both Pif1 and RecQ are difficult to identify.

We repeated the experiments to show the feasibility of the nanotensioner method. We observed distinct unwinding steps when the ATP concentration was reduced to 1 μM [Fig. 2(a). Additional examples are shown in Fig. S5]. We adopted an unbiased step-finding algorithm to find the steps (see details in the Supplemental Material [18], Fig. S6) [22]. The corresponding FRET values are in agreement with the theoretical calculations assuming that Pif1 unwinds 1 bp at a time. We built a histogram of the steps [Fig. 2(b)] in the FRET range from 0.2 to 0.8 in which an approximately linear relationship exists between the distance of the two dyes and the observed FRET value [46,47]. The distribution is narrow and has a peak at ΔFRET = 0.134 ± 0.001. The measurements at 0.5 μM ATP yielded similar results. We also built
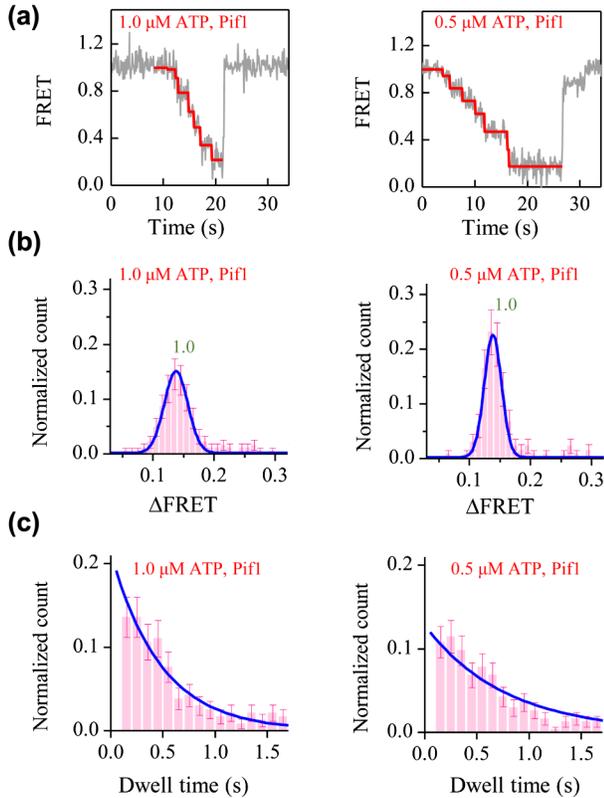
FIG. 2. Pif1 unwinds 1 bp at a time in the assay with nanotensioner. (a) Typical unwinding bursts with equidistant steps. The DNA rewinds as the enzyme disassociates. (b) Histograms of FRET jumps (ΔFRET). Blue lines represent Gaussian fittings with peak positions 0.134 ± 0.001 (SD = 0.03) (left) and 0.135 ± 0.001 (SD = 0.02) (right). The step sizes are indicated on top of each peak. (c) Histograms of dwell times before each stepping event. Blue lines represent single-exponential fittings with decay constants 0.44 ± 0.08 s (left) and 0.78 ± 0.06 s (right). The statistics are from > 150 traces at each ATP.

histograms of dwell times before each stepping event [Fig. 2(c)]. These distributions can be described by a single-exponential function, suggesting that these dwell times are governed by a single kinetic event. It is an expected behavior for binding of a single ATP molecule before each step under limiting ATP concentrations. The decay time (0.44 ± 0.08 s) at 1 μM ATP is about half of that (0.78 ± 0.06 s) at 0.5 μM ATP, suggesting that the dwell is dominated by the time the protein takes to bind a single ATP. Altogether, the results are consistent with a model in which the enzyme hydrolyzes a single ATP at a time and unzips the duplex with a uniform step size of 1 bp.

In contrast to the results of Pif1 in Fig. 2, the unwinding steps of RecQ are not equidistant even with the nanotensioner substrate. Shown in Fig. 3 are representative unwinding bursts of RecQ recorded at low ATP concentrations. We observed events in which the DNA overhangs increased in length with various increments. The distributions of the increments, i.e., the step sizes, are composed of multiple peaks. Besides the expected peak corresponding to
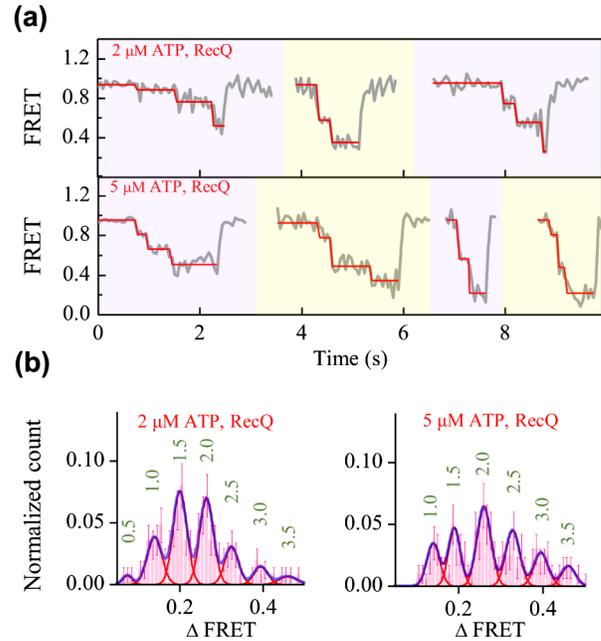

FIG. 3. Nonuniform stepping of RecQ in the assay with nanotensioner. (a) Typical unwinding bursts at 2 μM (top) and 5 μM (bottom) ATP. The individual traces are shifted horizontally for clarity. (b) The corresponding histograms of ΔFRET. Blue lines in the histograms represent Gaussian fittings. The step sizes converted from the peak positions are indicated on top of each peak. The data are from > 200 traces at each ATP.

unwinding of 1 bp, the distribution at 2 μM ATP shows other peaks corresponding to unwinding of 0.5, 1.5, 2.0, 2.5, and 3.0 bp, respectively. Similarly, the distribution of step sizes at 5 μM ATP also shows a few major peaks, centered at 1.0, 1.5, 2.0, 2.5, 3.0, and 3.5 bp, respectively.

It is obvious that the unwinding kinetics of RecQ is very different from that of Pif1. The difference is detectable only when the resolution is better than 0.5 bp. Even with a 1-bp resolution, one may still not be able to detect the 0.5-bp-increment steps and get only a smeared wide peak in the histogram, hence, mistakenly drawing a conclusion that RecQ unwinds DNA with step sizes of 2–4 bp. Irrespective of the big differences, we show in the following that one can have a general understanding of the stepping kinetics of Pif1, RecQ, and, possibly, other helicases. To this end, we argue that (i) a helicase breaks a base pair upon ATP hydrolysis, generating single-stranded nucleotides, 1 nt for each tail of the forked DNA, (ii) unwinding steps, however, are observable only when the generated nucleotides are released, and (iii) in general, it is not necessary that a nucleotide be released concomitantly with its generation because it may stick to some domain of the helicase through, e.g., electrostatic interaction and/or hydrogen bonds [10,48]. In other words, the helicase may sequester the nascent nucleotides and then release them after a random number of 1-bp opening events (Fig. 4). The very general depiction involves three independent parameters, i.e., a base-pair breaking rate k_b, a 3′-tail
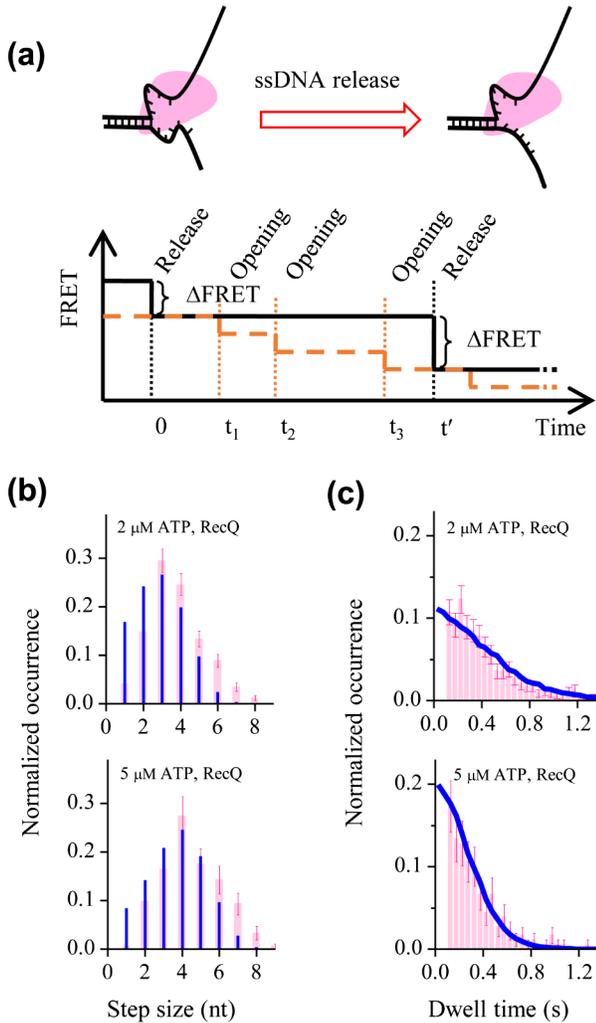
FIG. 4. Stepping kinetics of helicase-catalyzed DNA unwinding. (a) Model of helicase-catalyzed dsDNA unwinding and schematic time course of observable FRET signal (black line). At $t = 0$ and t', FRET jumps appear because of the release of the sequestered nascent nucleotides on the $3'$ or $5'$ strand. The orange dashed line represents the expected time course of the FRET signal if nucleotide release is synchronous with base pair opening, as for Pif1. In the time duration from $t = 0$ to t', base pair opening occurs three times at t_1, t_2, and t_3, respectively. (b),(c) Comparison between the Monte Carlo simulated (blue) and the experimentally measured (pink) probabilities of step sizes and dwell times of RecQ. The experimental probabilities in (b) are integrals of each Gaussian function in Fig. 3.

releasing rate k_{r1}, and a $5'$-tail releasing rate k_{r2}. The base-pair breaking rate k_b of a helicase is regulated by the concentration of ATP and can be derived from the unwinding rate versus ATP curve. There is, however, no direct way to calculate the two tail-releasing rates k_{r1} and k_{r2}. They can be estimated by using Monte Carlo simulations, detailed in the Supplemental Material [18], to reconstruct the histograms of dwell times and step sizes according to the kinetic model sketched in Fig. 4.

In our Monte Carlo simulations we made an assumption that the two tail-releasing rates, k_{r1} and k_{r2}, increase with

the number of nucleotides sequestered by the helicase. This is necessary because, otherwise, when k_b becomes much larger than k_{r1} and k_{r2} at high ATP concentrations, the number of nucleotides held by the helicase might become too long to be true. A few physical factors may underlie the assumption. Plausible ones include the reduction of entropy due to confinement of the nucleotides and/or the increase in energy due to accumulation of the negative charges of DNA on the surface of the protein. For simplicity, we assume that the extra energy is proportional to the number of nucleotides held by the helicase. The tail-releasing rates can hence be written as $k_{ri} = k_{ri}^0 \exp[\alpha(n - 1)]$, where $i = 1$ or 2, n is the number of nucleotides sequestered by the enzyme and k_{ri}^0 is the rate when $n = 1$. Using the values of the breaking rate k_b from literature [49], 3.3 s^{-1} at 2 μM ATP and 6.8 s^{-1} at 5 μM ATP, the reconstructed histograms for RecQ resemble the measured ones when the following parameters are used: $k_{r1}^0 \approx k_{r2}^0 = 0.3$ s^{-1} and $\alpha = 0.7$ (Figs. 4(c) and 4(d); Fig. S7 in the Supplemental Material [18]). We can also use the Monte Carlo simulation to reconstruct the histograms for Pif1 (Fig. S8 in the Supplemental Material [18]). The nucleotide releasing rates of Pif1 must be orders of magnitude higher than the base-pair breaking rate in order to have a good fit. This is equivalent to saying that the nucleotides are released immediately after they are generated. As a consequence, the observable unwinding step sizes is 1 bp and the dwell time distributions are exponential with characteristic dwell times depending on the concentration of ATP [Fig. 2(c)]. Taken together, the kinetic model sketched in Fig. 4 applies to both helicases with apparently different unwinding behaviors. In addition, a few more simulations with different parameters (Fig. S7 in the Supplemental Material [18]) implied that a histogram of dwell times does not necessarily follow a simple function; it may even not be monotonic under certain conditions.

smFRET has become the technique of choice to study helicases. However, to the best of our knowledge, it had not yet been able to resolve 1-nt step size before our work. On one hand, the unprecedented resolution provided by the presented nanotensioner approach enabled us to reveal the details of helicase-catalyzed DNA unwinding that are not easy to study with conventional smFRET method. On the other hand, the high-resolution data also allowed us to propose a unified molecular mechanism for the two helicases that belong to two different superfamilies with apparently different unwinding behaviors, implying that many helicases might be more fundamentally correlated. We speculate that the technique used in the current work will find wide applications in smFRET studies of many other molecular motors that convert dsDNA to ssDNA and vice versa such as helicases and polymerases.

[*]yinglu@iphy.ac.cn
[†]mingli@iphy.ac.cn

[1] S. S. Patel and K. M. Picha, Annu. Rev. Biochem. **69**, 651 (2000).
[2] S. S. Patel and I. Donmez, J. Biol. Chem. **281**, 18265 (2006).
[3] M. R. Singleton, M. S. Dillingham, and D. B. Wigley, Annu. Rev. Biochem. **76**, 23 (2007).
[4] T. M. Lohman, E. J. Tomko, and C. G. Wu, Nat. Rev. Mol. Cell Biol. **9**, 391 (2008).
[5] J. A. Ali and T. M. Lohman, Science **275**, 377 (1997).
[6] B. Sikora, R. L. Eoff, S. W. Matson, and K. D. Raney, J. Biol. Chem. **281**, 36110 (2006).
[7] R. Ramanagoudr-Bhojappa, S. Chib, A. K. Byrd, S. Aarattuthodiyil, M. Pandey, S. S. Patel, and K. D. Raney, J. Biol. Chem. **288**, 16185 (2013).
[8] Z. Qi, R. A. Pugh, M. Spies, and Y. R. Chemla, eLife **2**, e00334 (2013).
[9] S. Dumont, W. Cheng, V. Serebrov, R. K. Beran, I. Tinoco, Jr., A. M. Pyle, and C. Bustamante, Nature (London) **439**, 105 (2006).
[10] W. Cheng, S. G. Arunajadai, J. R. Moffitt, I. Tinoco, Jr., and C. Bustamante, Science **333**, 1746 (2011).
[11] I. De Vlaminck et al., Nano Lett. **11**, 5489 (2011).
[12] D. Dulin, T. J. Cui, J. Cnossen, M. W. Docter, J. Lipfert, and N. H. Dekker, Biophys. J. **109**, 2113 (2015).
[13] T. Ha, I. Rasnik, W. Cheng, H. P. Babcock, G. H. Gauss, T. M. Lohman, and S. Chu, Nature (London) **419**, 638 (2002).
[14] S. Syed, M. Pandey, S. S. Patel, and T. Ha, Cell Rep. **6**, 1037 (2014).
[15] S. Myong, M. M. Bruno, A. M. Pyle, and T. Ha, Science **317**, 513 (2007).
[16] G. Lee, M. A. Bratkowski, F. Ding, A. L. Ke, and T. Ha, Science **336**, 1726 (2012).
[17] M. C. Murphy, I. Rasnik, W. Cheng, T. M. Lohman, and T. J. Ha, Biophys. J. **86**, 2530 (2004).
[18] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.119.138102 for further details on sample preparation and data analyses, which includes Refs. [12, 19–22].
[19] J. B. Boule and V. A. Zakian, Nucleic Acids Res. **35**, 5809 (2007).
[20] H. Q. Xu, E. Deprez, A. H. Zhang, P. Tauc, M. M. Ladjimi, J. C. Brochon, C. Auclair, and X. G. Xi, J. Biol. Chem. **278**, 34925 (2003).
[21] S. B. Smith, Y. Cui, and C. Bustamante, Science **271**, 795 (1996).
[22] J. W. Kerssemakers, E. L. Munteanu, L. Laan, T. L. Noetzel, M. E. Janson, and M. Dogterom, Nature (London) **442**, 709 (2006).
[23] M. Lee, S. H. Kim, and S. C. Hong, Proc. Natl. Acad. Sci. U.S.A. **107**, 4985 (2010).
[24] X. Long, J. W. Parks, C. R. Bagshaw, and M. D. Stone, Nucleic Acids Res. **41**, 2746 (2013).
[25] F. E. Kemmerich, M. Swoboda, D. J. Kauert, M. S. Grieb, S. Hahn, F. W. Schwarz, R. Seidel, and M. Schlierf, Nano Lett. **16**, 381 (2016).
[26] M. J. Comstock, K. D. Whitley, H. Jia, J. Sokoloski, T. M. Lohman, T. Ha, and Y. R. Chemla, Science **348**, 352 (2015).
[27] Q. Guo, Y. He, and H. P. Lu, Phys. Chem. Chem. Phys. **16**, 13052 (2014).
[28] B. Choi, G. Zocchi, Y. Wu, S. Chan, and L. J. Perry, Phys. Rev. Lett. **95**, 078102 (2005).
[29] B. Choi, G. Zocchi, S. Canale, Y. Wu, S. Chan, and L. J. Perry, Phys. Rev. Lett. **94**, 038103 (2005).
[30] J. B. Bessler, J. Z. Torres, and V. A. Zakian, Trends Cell Biol. **11**, 60 (2001).
[31] M. L. Bochman, N. Sabouri, and V. A. Zakian, DNA Repair (Amst) **9**, 237 (2010).
[32] J. H. Li et al., Nucleic Acids Res. **44**, 4330 (2016).
[33] V. P. Schulz and V. A. Zakian, Cell **76**, 145 (1994).
[34] J. Zhou, E. K. Monson, S. C. Teng, V. P. Schulz, and V. A. Zakian, Science **289**, 771 (2000).
[35] S. Makovets and E. H. Blackburn, Nat. Cell Biol. **11**, 1383 (2009).
[36] K. Paeschke, J. A. Capra, and V. A. Zakian, Cell **145**, 678 (2011).
[37] P. Janscak, P. L. Garcia, F. Hamburger, Y. Makuta, K. Shiraishi, Y. Imai, H. Ikeda, and T. A. Bickle, J. Mol. Biol. **330**, 29 (2003).
[38] D. A. Bernstein and J. L. Keck, Nucleic Acids Res. **31**, 2778 (2003).
[39] M. E. Fairman-Williams, U. P. Guenther, and E. Jankowsky, Curr. Opin. Struct. Biol. **20**, 313 (2010).
[40] H. Q. Xu, E. Deprez, A. H. Zhang, P. Tauc, M. M. Ladjimi, J. C. Brochon, C. Auclair, and X. G. Xi, J. Biol. Chem. **278**, 34925 (2003).
[41] W. K. Chu and I. D. Hickson, Nat. Rev. Cancer **9**, 644 (2009).
[42] L. Bjergbaek, J. A. Cobb, and S. M. Gasser, Swiss Med. Wkly. **132**, 433 (2002).
[43] J. A. Cobb, L. Bjergbaek, and S. M. Gasser, FEBS Lett. **529**, 43 (2002).
[44] C. Bustamante, J. F. Marko, E. D. Siggia, and S. Smith, Science **265**, 1599 (1994).
[45] R. M. Harrison, F. Romano, T. E. Ouldridge, A. A. Louis, and J. P. K. Doye, arXiv:1506.09005v1.
[46] B. T. Harada, W. L. Hwang, S. Deindl, N. Chatterjee, B. Bartholomew, and X. Zhuang, eLife **5**, e10051 (2016).
[47] T. R. Blosser, J. G. Yang, M. D. Stone, G. J. Narlikar, and X. Zhuang, Nature (London) **462**, 1022 (2009).
[48] J. A. Newman et al., Nucleic Acids Res. **43**, 5221 (2015).
[49] D. Klaue, D. Kobbe, F. Kemmerich, A. Kozikowska, H. Puchta, and R. Seidel, Nat. Commun. **4**, 2024 (2013).