Communications

Real-Time Monitoring of DNA Polymerase Function and Stepwise Single-Nucleotide DNA Strand Translocation through a Protein Nanopore**

John Chu, Marcos González-López, Scott L. Cockroft, Manuel Amorín, and M. Reza Ghadiri

Single-molecule DNA sequencing with nanopores is a promising concept for advancing the next generation of rapid and low-cost sequencing technologies.[1,2] A proposed approach in this field involves moving a single-stranded DNA (ssDNA) through a narrow pore under the influence of an electric field and identifying the nucleobases as they pass by a reading-head located within the pore.[1–3] A key challenge is to work out how to ratchet DNA strand through a nanopore at the appropriate speed and time intervals that are commensurate with real-time sequential nucleobase identification.[1] Here we describe an approach that exploits the intrinsic motor function of DNA polymerases to effect stepwise DNA strand translocation through an alpha-hemolysin (αHL) protein pore. The system is based on a single-species supramolecular nanopore construct,[4] whose configurational dynamic features provide the means for real-time monitoring of DNA strand translocation and DNA polymerase function.

Any effort directed at reducing the nanopore DNA sequencing concept to practice, whether employing protein or solid-state pores,[5,6] must confront the key issues of sequential nucleobase recognition and controlling the speed and motion of ssDNA through the pore.[1] In regard to the nucleobase recognition issue, there have been promising recent advances in reading-head engineering with protein nanopores that can be used to discern the identity of a nucleotide at a specific location on a ssDNA based on the observed ion current levels if the ssDNA is mechanically immobilized inside the nanopore.[7,8] On the other hand, little progress has been made in controlling the translocation speed of ssDNA through nanopores.[9] The difficulty lies in the fact that, under the influence of an electric field, DNA translocation through a nanopore is too fast (1 to 20 µs per base for αHL) to allow base identification by standard single-channel

recording.[3,10] To overcome this obstacle, a method needs to be devised that can effect unidirectional movement of DNA through the nanopore in single nucleotide steps, and at an appropriate speed (1 to 100 ms per base) to provide sufficient residence time for resolving the ion current level associated with each nucleotide step.

We present here an approach that features a previously overlooked configurational dynamics of single-species αHL·DNA–PEG^{phos} rotaxanes (PEG^{phos} = a hexaethylene phosphate polymer, Figure 1).[4,11] Rotaxanes are supramolecular constructs consisting of a molecular thread mechanically interlocked within a ring (macrocycle) by bulky terminal substituents.[12] αHL·DNA–PEG^{phos} rotaxanes can adopt distinct configurations defined by the relative position and orientation of the thread to the ring.[11] The constructs employed in this study were prepared by threading a single αHL pore from the *trans* side with a 5′-streptavidin-capped biotinylated-DNA-PEG^{phos} strand and capturing its protruding 3′-tail on the *cis* side by a complementary DNA primer as described previously (Supporting Information, Figure S1).[4] It should be noted that although the appended PEG^{phos} segment is immaterial to the configurational dynamics of the rotaxanes (see below), it is nevertheless a useful design feature that facilitates structural and functional characterizations. The PEG^{phos} segment, by virtue of its distinct backbone charge distribution and steric demand, can give rise to characteristic ion current levels that are correlated with its relative position within the pore (Figure S2).[4] The polyanionic nature of the DNA–PEG^{phos} hybrid strand allows a voltage-dependent force to be exerted on the thread held inside αHL to drive its translocation in the direction of the applied electrical field.[13]

We hypothesized that under positive potentials (*cis* side defined as the ground) the rotaxane can adopt one of the two limiting configurations ***A*** or ***B*** depending on the magnitude of the applied electric field (Figure 1a). At relatively low potentials (low force exerted on the thread), the more energetically stable configuration ***A*** is expected to predominate in which the dsDNA segment remains on the outside butting against the *cis* opening of αHL. On the other hand, the energetically less stable configuration ***B*** should predominate at relatively high potentials where the larger electric force exerted on the thread serves to overcome the unfavorable energetic requirements for dsDNA entry, translocation, and confinement inside the αHL vestibule. A rotaxane in configuration ***B***, with its bulky dsDNA segment lodged in αHL, should have a higher barrier to ion flow and thus exhibit lower conductance as compared to configuration ***A***. Further-

[*] J. Chu, Dr. M. González-López, Dr. S. L. Cockroft, Dr. M. Amorín, Prof. M. R. Ghadiri
Departments of Chemistry and Molecular Biology and
The Skaggs Institute of Chemical Biology
The Scripps Research Institute
10550 North Torrey Pines Road, La Jolla, CA 92037 (USA)
Fax: (+1) 858-784-2798
E-mail: ghadiri@scripps.edu
Homepage: http://www.scripps.edu/chem/ghadiri
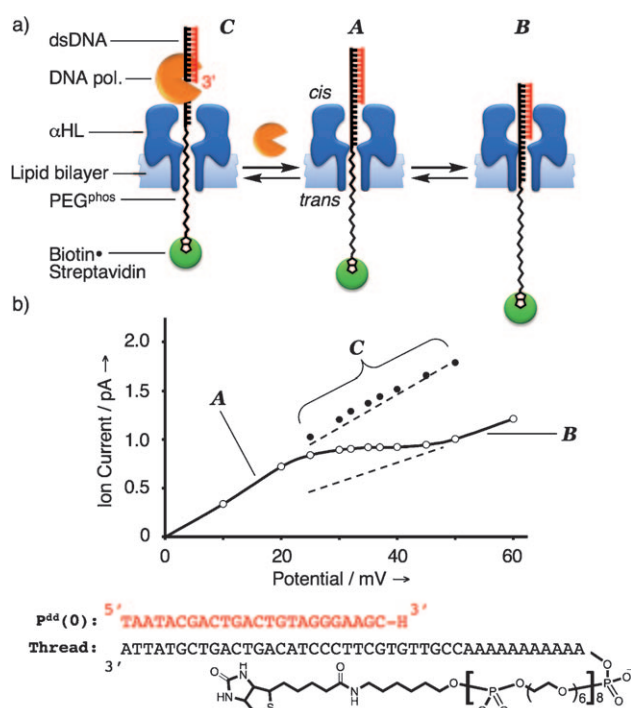
Figure 1. a) The free rotaxane is hypothesized to adopt two limiting configurations **A** and **B** depending on the force exerted on the thread by the transmembrane potential. In configuration **A**, the dsDNA segment is butting against the *cis* opening of the αHL protein pore. The rotaxane assumes configuration **B** when a greater electric force is applied to overcome the energetic requirements for dsDNA entry and confinement within the αHL vestibule. The 3′-terminus of the DNA primer in the context of the rotaxane is only accessible in configuration **A** to DNA polymerases in the *cis* chamber. The rotaxane adopts configuration **C** upon DNA polymerase binding. b) A rotaxane was constructed using αHL, thread, and DNA primer **P**^{dd}**(0)** for this study. Its measured *I–V* curve displays a pronounced deflection between +25 and +50 mV (*cis* side defined as ground). The linear regions at relatively low and high applied potentials correspond to the rotaxane in configurations **A** and **B**, respectively, whereas the rotaxane in the deflection region is assumed to be metastable and undergo rapid dynamic exchange between the two configurations. Interestingly, the deflection manifests itself as the moving average of regions **A** and **B** (linear extrapolations are shown as dashed lines). Each data point in the *I–V* curve (open circles) is the average of six measurements; the solid data points represent ion current levels of the rotaxane in the KF(exo⁻) bound configuration **C** under various applied potentials (see Figure S5). Error bars are smaller than the diameter of the data points.

more, because there are multiple and increasingly higher energetic structures in progressing from configuration **A** to **B**, the current–voltage (*I–V*) curve spanning this transition should be relatively broad and likely reflect the moving average of **A** and **B** conductance levels. Consistent with this hypothesis, pronounced deflections (typically spanning about 25 mV) were observed in the recorded *I–V* curves of αHL·DNA–PEG^{phos} rotaxanes in the region between +25 and +75 mV depending on the length of the DNA primer employed (Figure 1 b and Figure S2). Within this region of the *I–V* curve the rotaxane is considered to be configurationally metastable and fluctuating rapidly between **A** and **B** (Figure S3). Therefore, we reasoned that as we seek to harness the intrinsic motor function of DNA polymerases to effect

stepwise DNA strand translocation through αHL,[4,14] such configurational dynamics should provide the means for the hitherto inaccessible real-time monitoring of the DNA strand translocation process. DNA polymerases recognize and bind to the 3′ segment of a DNA primer–template duplex (P/T). In the context of the rotaxane system employed here, this binding site is accessible only from configuration **A** because the 3′ terminus of the primer is shielded within the αHL vestibule in configuration **B**. Consequently, DNA polymerase binding to the rotaxane should change the configurational dynamics away from the fluctuation between **A** and **B**. Furthermore, based on structural considerations, the resulting polymerase·αHL·DNA–PEG^{phos} rotaxane complex now residing in the new configuration **C** should give a distinctly higher ion conductance level than the moving average of **A** and **B** (Figure 1 b).

To test our hypothesis, we prepared a rotaxane by threading a DNA–PEG^{phos} hybrid strand through the αHL and completed the interlocked structure using a 24-nucleotide 2′,3′-dideoxy-terminated DNA primer, **P**^{dd}**(0)**. The measured *I–V* curve bore the expected deflection between +25 and +50 mV, which is indicative of rapid fluctuation between configurations **A** and **B** within this potential region (Figure 1 b). Next, KF(exo⁻) (3′→5′ exonuclease deficient Klenow fragment of *E. coli* DNA polymerase I) was introduced into the *cis* chamber and the ion current was recorded continuously at +35 mV. Transient spikes at a higher ion current level were observed shortly after the addition of KF(exo⁻) (Figure 2 a). These events are consistent with KF(exo⁻) interacting with the rotaxane P/T duplex shifting the rotaxane to configuration **C**, and resulting in a transient higher ion conductance level. To confirm that the higher ion current events correspond to repeated association and dissociation of KF(exo⁻) to the αHL·DNA–PEG^{phos} rotaxane, an excess amount of free DNA P/T duplex was added to the *cis* chamber to compete with the rotaxane for KF(exo⁻) binding. Indeed, all ion current spikes vanished following the addition of P/T (Figure S4). Furthermore, ion current recordings at various applied potentials spanning the deflection region of the *I–V* curve established that both the frequency and lifetimes of KF(exo⁻) binding events decrease with increasing applied potential (Figure 1 b and Figure S5). These observed trends are consistent with our working hypothesis that as the applied potential is increased, a larger electric force is exerted on the thread and consequently there is a greater probability for rotaxane to reside in configuration **B** which together give rise to lower observed frequency of KF(exo⁻) binding events. Moreover, the thread is pulled more strongly into the αHL vestibule at higher potentials, which counteracts the free energy of KF(exo⁻) association, and thus the decreased lifetimes of KF(exo⁻) binding events. The above phenomena seem to be general and were also observed using a rotaxane constructed in the opposite configuration with the dsDNA segment and the biotin–streptavidin complex on the *trans* and *cis* side, respectively (Figures S1 and S6), other types of DNA polymerases (9°N$_m$ and the *B. strearothermophilus* (*Bst*) large fragment), and single-species rotaxanes that were assembled from αHL and DNA only threads (Figure S7).
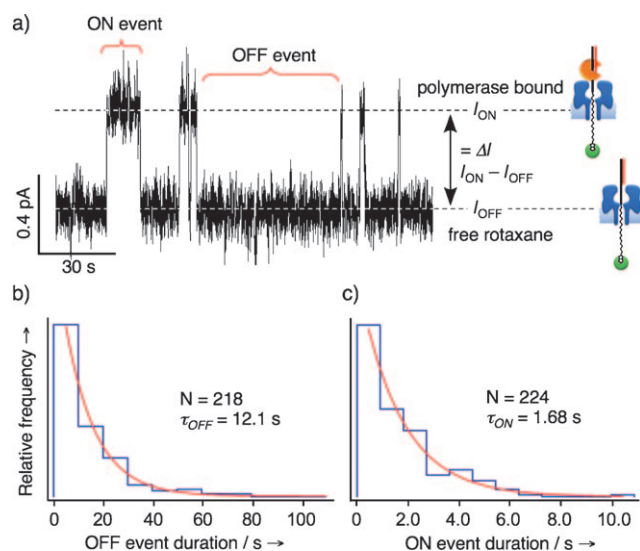
Figure 2. a) Multiple KF(exo⁻) binding events to the rotaxane (Figure 1b) is evident in the section of the ion current trace recorded at +35 mV shown here. The baseline level (I_{OFF}) and elevated (I_{ON}) ion current levels correspond to the free rotaxane and the DNA polymerase bound states, respectively. b,c) The ON and OFF event durations are plotted as histograms ($N = 224$ and 218) with bin widths determined by the Scott formula,[15] which are 9.90 and 0.91 s, respectively. The histograms were each fitted to a single exponential decay function to obtain the kinetic constants for the dissociation and association processes ($\tau_{OFF} = 12.1$ s and $\tau_{ON} = 1.68$ s, see Supporting Information). This translates to a dissociation rate constant (k_{OFF}) of 0.6 s⁻¹ and an association rate constant (k_{ON}) of $1.1 \times 10^7 \, \text{M}^{-1} \text{s}^{-1}$. The observed k_{ON} is similar to the reported literature values derived from ensemble measurements, whereas comparatively the k_{OFF} is about an order of magnitude larger[16]—likely due to the applied force condition of our study.[13]

We next examined the utility of the single-species αHL·DNA–PEG^phos rotaxane constructs for real-time monitoring of KF(exo⁻)-catalyzed DNA primer extensions. In this regard, the PEG^phos segment is an important design feature that enables facile monitoring of single nucleotide incorporation events as its relative position within the αHL pore can be identified through characteristic changes in the observed ion conductance levels.[4] Since αHL is too narrow for dsDNA to pass through, primer extension in effect moves the DNA–PEG^phos thread one base at a time in the *trans* to *cis* direction, which result in pulling a corresponding PEG^phos segment into the pore (Figure S2). Because of the smaller steric demand of the PEG^phos segment as compared to DNA, higher PEG^phos to DNA ratio within the αHL pore blocks the nanopore to a lesser extent and allows higher ion current flow, thereby establishing a correlation between primer length and ion current. We have reported previously that such correlation can be used to sensitively and consistently distinguish single-nucleotide primer length differences.[4] To assess the effectiveness of this process for real-time monitoring of DNA polymerase function, a single-species rotaxane was assembled using αHL, the DNA–PEG^phos hybrid thread, and a 24-nucleotide DNA primer **P(0)** strand bearing the 3′-OH group necessary for DNA polymerase catalyzed strand extension. The primer extension reactions were performed at a constant

applied potential of +35 mV, indicated by the aforementioned studies to be the optimal applied potential for real-time monitoring of DNA polymerase function. KF(exo⁻) was first added to the *cis* side of the rotaxane in order to establish a working system signified by the transient high ion conductance spikes resulting from reversible association and dissociation of the polymerase to the dsDNA segment of the rotaxane. Next, the required deoxynucleotide triphosphates (dATP, dCTP, dGTP) were introduced to the *cis* chamber to initiate primer extension reactions. It should be noted that dTTP was excluded from the reaction mixture to limit the total number of DNA polymerase catalyzed primer extension steps and provide a defined reaction end point. Figure 3
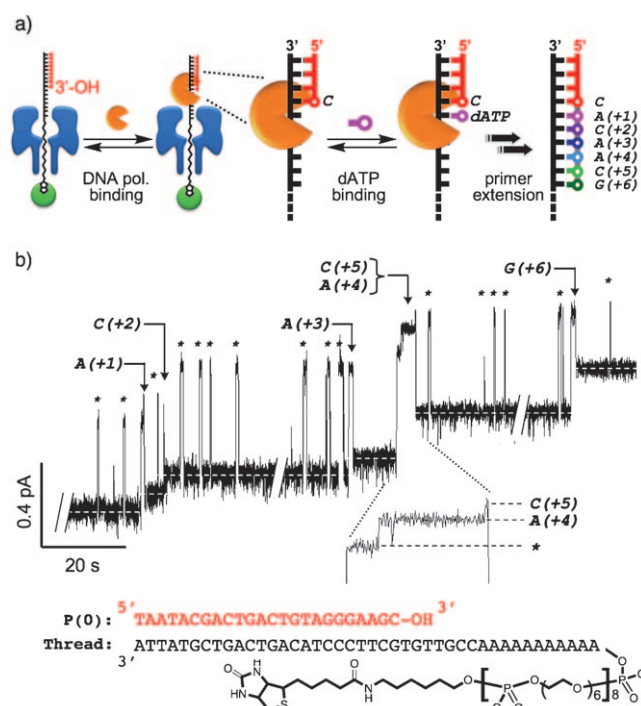
Figure 3. a) KF(exo⁻) can freely associate and dissociate to the dsDNA segment of the rotaxane. The DNA polymerase bound complex can recognize a nucleotide triphosphate (dNTP) to form the ternary complex. The primer is then extended by DNA polymerase catalyzed phosphodiester bond formation upon correct nucleobase pairing. b) Reversible non-productive DNA polymerase bindings to the dsDNA segment of the rotaxane are manifest as spikes (marked by asterisks), whereas nucleotide incorporation events are marked by the irreversible stepwise baseline changes to increasingly higher ion conductance levels. Incorporation events are denoted as **N**(+**X**), where **N** is the inferred identity of the incorporated nucleotide and (+**X**) is the number of added nucleotides relative to the starting 24-nt primer. Low dNTP concentrations (20 nM of each dA, dC and dGTP) were employed in this experiment to favor a distributive DNA polymerase catalyzed nucleotide incorporation process. In all but one event the primer was extended in single nucleotide steps. The one exception is the processive incorporation of **A**(+4) and **C**(+5), in which two steps are clearly resolved in the KF(exo⁻) bound state (inset). The thread and DNA primer **P(0)** bearing the free 3′-OH group used to assemble the rotaxane are shown at the bottom. The ion current recordings are shown at 50 Hz resolution. See Figures S8–S11 for additional recordings, details regarding experimental conditions, data processing, and sample rate versus spatial resolution.

illustrates one such study where seven clearly resolved irreversible levels are evident in the ion current trace as a result of KF(exo⁻)-catalyzed stepwise primer extension events. Among the productive KF(exo⁻) binding events, one was a processive event where two nucleotides were incorporated, while the other four were distributive events where only a single nucleotide was incorporated before KF(exo⁻) dissociation (Figure 3b). The likely generality of this approach is supported by several real-time primer extension recordings using KF(exo⁻), KF(exo⁺), and *E. coli* DNA polymerase I. We also established that real-time monitoring of DNA polymerase function was also possible using a rotaxane of the opposite configuration, albeit with moderate signal to noise ratio (for these additional experiments see Figures S8–S10).

To estimate the spatial resolution of the nanopore device, we assumed that each nucleotide incorporation step shifted the thread by approximately 3.4 Å (base stacking distance in B-form DNA). The distance-to-current conversion scale of $\Delta L/\Delta I \approx 22.7$ Å pA⁻¹ was derived by averaging the observed ion current increments across multiple experiments ($\Delta I \approx 150$ fA for each primer extension step,[4] see also Figures S8 and S9). Although noise is intrinsic to any single-molecule measurement, the signal to noise ratio can be enhanced by averaging the crude data at the expense of temporal resolution. Based on the estimated $\Delta L/\Delta I$ conversion scale, the highest sample rate capable of achieving single-base resolution at 95 % confidence with our current setup is approximately 50 Hz (Figure S11), representing a combined temporal and spatial resolution that bodes well among the current state of the art single-molecule techniques.[17]

In summary, we have described the design, mechanistic rationale, and functional characterization of a supramolecular nanopore device that harnesses the motor function of DNA polymerases to specify both the speed and ratcheted (base by base) motion of ssDNA translocation through a protein nanopore. Importantly, the inherent configurational dynamics of this system allows single-molecule DNA polymerase binding and function, as well as successive single-nucleotide DNA translocation steps to be monitored in real-time with high temporal and spatial resolution. The likely prospect of fabricating an integrated nanopore device that couples the design features reported here with the recent advances in reading head engineering with αHL,[7] suggests a plausible path toward realizing real-time single-molecule nanopore DNA sequencing.

[1] D. Branton et al., *Nat. Biotechnol.* **2008**, *26*, 1146 (see Supporting Information for full reference).
[2] H. Bayley, *Curr. Opin. Chem. Biol.* **2006**, *10*, 628.
[3] J. J. Kasianowicz, E. Brandin, D. Branton, D. W. Deamer, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 13770.
[4] a) S. L. Cockroft, J. Chu, M. Amorin, M. R. Ghadiri, *J. Am. Chem. Soc.* **2008**, *130*, 818; b) F. Olasagasti, K. R. Lieberman, S. Benner, G. M. Cherf, J. M. Dahl, D. W. Deamer, M. Akeson, *Nat. Nanotechnol.* **2010**, *5*, 798.
[5] a) J. Li, D. Stein, C. McMullan, D. Branton, M. J. Aziz, J. A. Golovchenko, *Nature* **2001**, *412*, 166; b) C. Dekker, *Nat. Nanotechnol.* **2007**, *2*, 209.
[6] a) T. Z. Butler, M. Pavlenok, I. M. Derrington, M. Niederweis, J. H. Gundlach, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 20647; b) J. Clarke, H. C. Wu, L. Jayasinghe, A. Patel, S. Reid, H. Bayley, *Nat. Nanotechnol.* **2009**, *4*, 265.
[7] a) D. Stoddart, A. J. Heron, E. Mikhailova, G. Maglia, H. Bayley, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 7702; b) D. Stoddart, G. Maglia, E. Mikhailova, A. J. Heron, H. Bayley, *Angew. Chem.* **2010**, *122*, 566; *Angew. Chem. Int. Ed.* **2010**, *49*, 556; c) D. Stoddart, A. J. Heron, J. Klingelhoefer, E. Mikhailova, G. Maglia, H. Bayley, *Nano Lett.* **2010**, *10*, 3633.
[8] N. Ashkenasy, J. Sanchez-Quesada, H. Bayley, M. R. Ghadiri, *Angew. Chem.* **2005**, *117*, 1425; *Angew. Chem. Int. Ed.* **2005**, *44*, 1401.
[9] a) A. Meller, L. Nivon, E. Brandin, J. Golovchenko, D. Branton, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 1079; b) D. Fologea, J. Uplinger, B. Thomas, D. S. McNabb, J. Li, *Nano Lett.* **2005**, *5*, 1734; c) E. H. Trepagnier, A. Radenovic, D. Sivak, P. Geissler, J. Liphardt, *Nano Lett.* **2007**, *7*, 2824; d) G. Maglia, M. R. Restrepo, E. Mikhailova, H. Bayley, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 19720.
[10] A. Meller, L. Nivon, D. Branton, *Phys. Rev. Lett.* **2001**, *86*, 3435.
[11] J. Sánchez-Quesada, A. Saghatelian, S. Cheley, H. Bayley, M. R. Ghadiri, *Angew. Chem.* **2004**, *116*, 3125; *Angew. Chem. Int. Ed.* **2004**, *43*, 3063.
[12] V. V. Balzani, A. Credi, F. M. Raymo, J. F. Stoddart, *Angew. Chem.* **2000**, *112*, 3484; *Angew. Chem. Int. Ed.* **2000**, *39*, 3348.
[13] U. F. Keyser, B. N. Koeleman, S. Van Dorp, D. Krapf, R. M. M. Smeets, S. G. Lemay, N. H. Dekker, C. Dekker, *Nat. Phys.* **2006**, *2*, 473.
[14] S. Benner, R. J. A. Chen, N. A. Wilson, R. Abu-Shumays, N. Hurt, K. R. Lieberman, D. W. Deamer, W. B. Dunbar, M. Akeson, *Nat. Nanotechnol.* **2007**, *2*, 718.
[15] D. W. Scott, *Biometrika* **1979**, *66*, 605.
[16] a) R. D. Kuchta, V. Mizrahi, P. A. Benkovic, K. A. Johnson, S. J. Benkovic, *Biochemistry* **1987**, *26*, 8410; b) H. Matsuno, K. Niikura, Y. Okahata, *Chem. Eur. J.* **2001**, *7*, 3305.
[17] a) B. Maier, D. Bensimon, V. Croquette, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 12002; b) W. J. Greenleaf, S. M. Block, *Science* **2006**, *313*, 801; c) J. R. Moffitt, Y. R. Chemla, D. Izhaky, C. Bustamante, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 9006; d) T. D. Harris et al., *Science* **2008**, *320*, 106; e) J. Eid et al., *Science* **2009**, *323*, 133 (see Supporting Information for full references); f) B. Gyarfas, F. Olasagasti, S. Benner, D. Garalde, K. R. Lieberman, M. Akeson, *ACS Nano* **2009**, *3*, 1457.