Single-molecule mechanical identification and sequencing

Fangyuan Ding[1,2], Maria Manosas[1,3], Michelle M Spiering[4], Stephen J Benkovic[4], David Bensimon[1,2,5], Jean-François Allemand[1,2] & Vincent Croquette[1,2]

High-throughput, low-cost DNA sequencing has emerged as one of the challenges of the postgenomic era. Here we present the proof of concept for a single-molecule platform that allows DNA identification and sequencing. In contrast to most present methods, our scheme is not based on the detection of the fluorescent nucleotides but on DNA hairpin length. By pulling on magnetic beads tethered by a DNA hairpin to the surface, the molecule can be unzipped. In this open state it can hybridize with complementary oligonucleotides, which transiently block the hairpin rezipping when the pulling force is reduced. By measuring from the surface to the bead of a blocked hairpin, one can determine the position of the hybrid along the molecule with nearly single-base precision. Our approach can be used to identify a DNA fragment of known sequence in a mix of various fragments and to sequence an unknown DNA fragment by hybridization or ligation.

Chain-terminator Sanger sequencing has dominated the DNA sequencing field for almost 20 years[1]. The need for faster and cheaper methods has driven the development of other approaches. These 'next'-generation DNA sequencing platforms[2–10] achieve high throughput by monitoring in parallel the successive incorporation of fluorescently labeled nucleotides by DNA polymerase or ligase in a very large number of microscopic vessels, each containing thousands of PCR-amplified copies of a short DNA fragment. However, owing to the limited read length as well as the complexity and bias of the required pre-amplification step, so-called 'third'-generation sequencing platforms have been developed[11–13]. By directly monitoring the incorporation of fluorescently labeled nucleotides in an array of single DNA molecules, the latest platforms do away with preamplification and allow for longer read length. However, these single-molecule sequencing methods are still plagued by the high cost of the labeled nucleotides and have high error rates (4–15%) because of low signal-to-noise ratios and non- or misdetection of the fluorescent signal[11]. Although nonfluorescent single-molecule sequencing alternatives have been proposed[14–16], they are not yet competitive with the fluorescence-based methods.

In parallel with new sequencing technologies, scientists have developed high-throughput methods for large-scale genome analyses, such as gene identification, single-nucleotide polymorphism detection and gene expression profiling, in particular cDNA library characterization[17]. To identify and quantify DNA fragments of known sequence in a given sample, DNA microarrays are often used[18,19]. However, this approach suffers from the need to preamplify the target DNA, and resulting data are limited by nonspecific hybridization, adsorption and the need for quantification of fluorescence[18,19].

Here we present a proof of concept of single-molecule identification and sequencing methods that do not rely on fluorescence but on the measurement of the extension of a DNA hairpin, that is, the distance between one end anchored to a surface and the other bound to a magnetic bead pulled away by small magnets (**Fig. 1**).

RESULTS

Detection of roadblocks in the rezipping pathway of a hairpin

We attached a DNA hairpin at one end to a coverslip via a digoxigenin–anti-digoxigenin bond and at the other to a magnetic bead via a biotin-streptavidin bond. This DNA hairpin can be generated in various ways. For example, it can be formed by ligation of a genomic DNA fragment to a DNA loop at one end and to a DNA fork structure labeled with biotin and digoxigenin at its other end (**Fig. 1a** and **Supplementary Fig. 1**).

Small magnets above the sample apply a vertical force on the tethered beads. The end-to-end distance of the hairpin is deduced in real time from the bead's image[20]. Although the setup is similar to the DNA unzipping experiment performed with optical tweezers[21], the use of a magnetic trap allows for a high degree of parallelism through the simultaneous application of the same force on many molecules[20,22].

We modulated the force to periodically open and close the DNA hairpins in a solution[23] containing oligonucleotides complementary to a section of the hairpin. For instance, an 83 base-pair (bp) hairpin periodically unfolded upon application of a force, F_{open}, (>15 pN) and rezipped upon a reduction in the force to F_{test}

[1]Laboratoire de Physique Statistique, Ecole Normale Supérieure, Université Pierre et Marie Curie Université Paris 06, Université Paris Diderot, Centre National de la Recherche Scientifique, Paris, France. [2]Institut de Biologie de l'Ecole Normale Supérieure, Paris, France. [3]Departament de Física Fonamental, Facultat de Física, Universitat de Barcelona, Barcelona, Spain. [4]Department of Chemistry, The Pennsylvania State University, University Park, Pennsylvania, USA. [5]Department of Chemistry and Biochemistry, University of California Los Angeles, Los Angeles, California, USA. Correspondence should be addressed to V.C. (vincent.croquette@lps.ens.fr).

RECEIVED 11 JULY 2011; ACCEPTED 17 JANUARY 2012; PUBLISHED ONLINE 11 MARCH 2012; DOI:10.1038/NMETH.1925

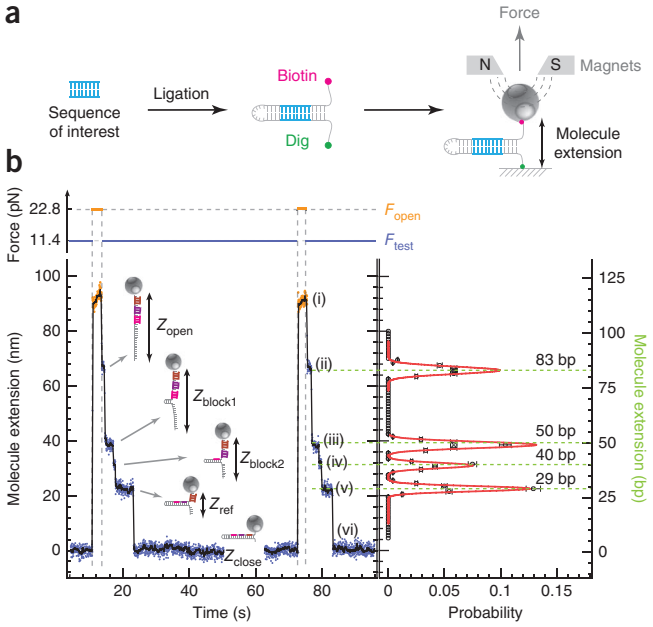

Figure 1 | Detection of oligonucleotide-induced blockages during rehybridization. (**a**) Hairpin construction design with target (sequence of interest) in the stem. Dig, digoxigenin; N and S represent magnet north and south poles. (**b**) Example of roadblocks resulting from the hybridization of three oligonucleotides on an 83-bp hairpin (Online Methods). Experimental traces recorded at F_{open} = 22.8 pN and F_{test} = 11.4 pN (left). Six different extension levels were observed: open hairpin at F_{open} (i), open hairpin at F_{test} (ii), partially annealed hairpin blocked by first (iii), second (iv) and third (v) oligos, and the folded hairpin (vi). The black curve corresponds to a 1 s average of the raw data. The black curve represents the histogram of the number of blockages ii–v per cycle at a given extension of the hairpin upon rehybridization at F_{test}: $Z = Z_{block} - Z_{close}$ in base pairs was obtained from ~77 force cycles on a single hairpin (right). Error bars were computed using the inverse of the square root of the number of blockages for each individual bin (1 nm). Gaussian fits to the data are shown in red, where the χ^2/n = 1.01 for blockage ii, 0.50 for iii, 0.732 for iv and 0.58 for v. Dashed green lines indicate the expected positions.

(~10 pN) (**Fig. 1b**). In the unfolded or open state, three different oligonucleotides in solution can hybridize to their individual complementary sequences on the hairpin. They transiently block the refolding of the hairpin at low force, which can be readily observed as three pauses in the time course of the hairpin's end-to-end distance measurements.

This measurement scheme provides the position and lifetime of the blockage along the hairpin. The opening of one base pair results in a change in the hairpin's end-to-end distance (extension), α_{ss}, of ~0.80 nm. With the current resolution of our apparatus (~1 nm), we can thus record the position of a blockage with an accuracy of about one nucleotide. The precise value of the applied force is not critical as long as it remains constant (**Supplementary Figs. 2–4** and **Supplementary Discussion**). Blockage lifetime, which is related to the stability of the hybrid, depended on the applied tension, the size of the complementary oligonucleotide, and the presence and location of mismatches between the oligonucleotide and the hairpin (**Supplementary Fig. 5**).

Identification by hybridization or single-cycle ligation

The identification of a desired DNA in a given sample is a relevant issue in many situations, for example, the detection of genetic mutations. Our scheme can be used for such identification. It requires the detection of a given hybridization fingerprint obtained with a set of short (~10-nucleotide (nt)) oligonucleotides chosen to hybridize perfectly along a particular DNA hairpin sequence. The fingerprint is obtained by testing for the existence of blockages during rezipping of the hairpins in the presence of probe oligonucleotide(s). We identified three DNA molecules with different sequences (1,241-bp, 83-bp and 179-bp; Online Methods) by hybridization of two different oligonucleotides of 10 nt and 11 nt (**Fig. 2**). Because we can detect the exact position of the hybrid,

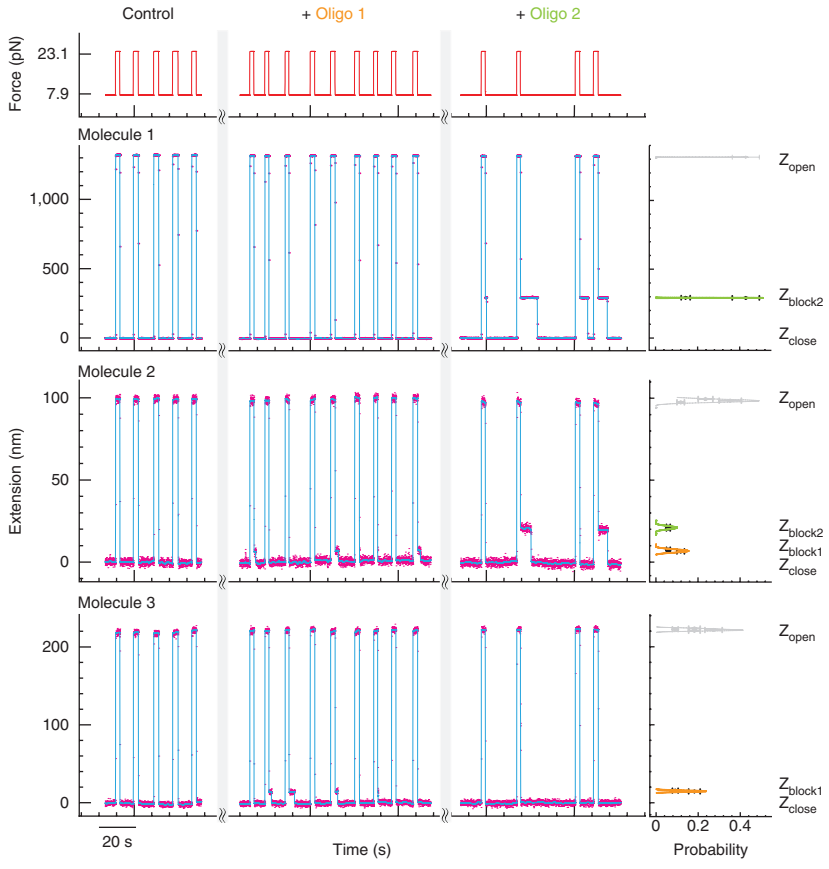

Figure 2 | Sequence identification by hybridization. Three hairpin-containing molecules (1, 2 and 3) were identified by oligonucleotides 1 and 2 (Online Methods). Experimental traces recorded during hairpin open-close cycles (left). Oligos 1 and 2 were added sequentially into the solution. Histogram of blockages from oligos 1 and 2 for each molecule (right); gray shading indicates full extension of the molecule. The error bars were computed using the inverse of the square root of the number of blockages for each individual bin (1 nm) ($n > 25$).

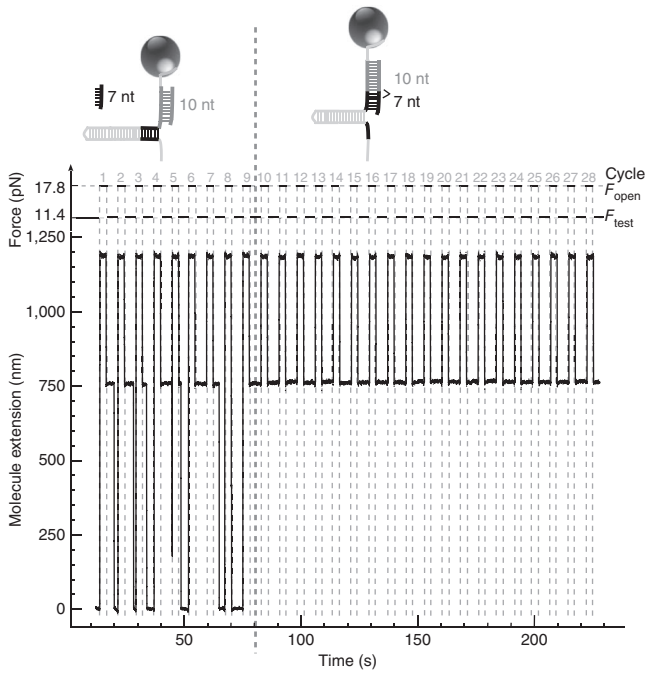

Figure 3 | Sequence identification by single ligation cycle. By applying a force F_{open} (>15 pN) a 1,241-bp DNA hairpin was opened periodically for 2 s to allow for the ligation between adjacent 7-nt and 10-nt oligomers hybridized to the hairpin stem (Online Methods). Upon ligation occurring at the 9th cycle, a quasi-permanent blockage, owing to the generated 17-nt oligomer, was observed from the 10th cycle to the 28th cycle.

it is not necessary to test each probe sequentially. Rather one may identify the blockage positions along the hairpin produced by the hybridization of the selected probes and use this hybridization pattern, instead of the mere presence of hybrids, as a fingerprint of the DNA sequence.

An alternative approach was to use the ligation of two adjacent short oligonucleotides (7 nt and 10 nt; **Fig. 3**) to target DNA as an identification tag. In the unzipped state of the hairpin, both sequences can hybridize to their adjacent target and be ligated. The blockage time observed during rezipping of the ligated 17-nt fragment was orders of magnitude longer than that of the unligated fragments, which was dominated by the blocking of the 10-nt oligo (a few seconds at ~10 pN), about 64 times longer than that of 7-nt oligo (**Fig. 3**). When we reduced the pulling force below 1 pN, the ligated fragment was expelled after ~1 min, thus allowing multiple verifications of identity.

Single-molecule sequencing by hybridization

Sequencing by hybridization has been demonstrated recently on an *Escherichia coli* genome cut into small fragments and amplified by a rolling-circle method. Using a complete set of 512 fluorescently labeled oligonucleotide 'tiles'[6] with five discriminating nucleotides, the authors tested for the presence of a subset of these tiles in each of the fragments. They then reconstructed the sequence by the overlapping fragments. As this approach only requires the detection of hybridization, it is compatible with our fluorescence-free platform, which provides additional information on the hybridization position. Using this extra information, we should be able to proceed in a similar fashion with a smaller number of discriminating nucleotides, thus reducing the number of hybridization tests. However, the manual buffer exchanges needed with our approach did not permit this high-throughput strategy, and we demonstrate sequencing by hybridization using only four tiles on a 31-bp DNA fragment.

In a process developed in the context of nanopore sequencing, DNA can be converted into a new sequence[24] in which each original nucleotide is encoded by a specific 8-nt sequence (A_8, T_8, G_8 and C_8) (**Fig. 4a**). The original DNA sequence can then be determined using our platform by sequential hybridization with four complementary oligomers. We implemented this scheme by recoding a 31-nt DNA fragment into a 248-bp dsDNA ligated with appropriate fragments to form a hairpin. Then we determined the sequence by hairpin opening and closing cycles done with oligonucleotides A_8, T_8, G_8 and C_8 successively added in the solution. As expected, the histogram of the blocking positions revealed the location of each expanded base (**Fig. 4b**). These positions along the recoded hairpin need to be determined with an accuracy of only 7 nm, well within the characteristics of our current setup.

Single-molecule sequencing by periodic ligation cycles

Sequencing DNA via successive ligation reactions is currently used in some 'next'-generation DNA sequencing schemes[25]. We implemented a similar protocol with our setup using the detection of the ligation of a complementary oligonucleotide to a growing primer to determine the underlying sequence. We tested the ligation of a complete set of oligonucleotides to a growing DNA

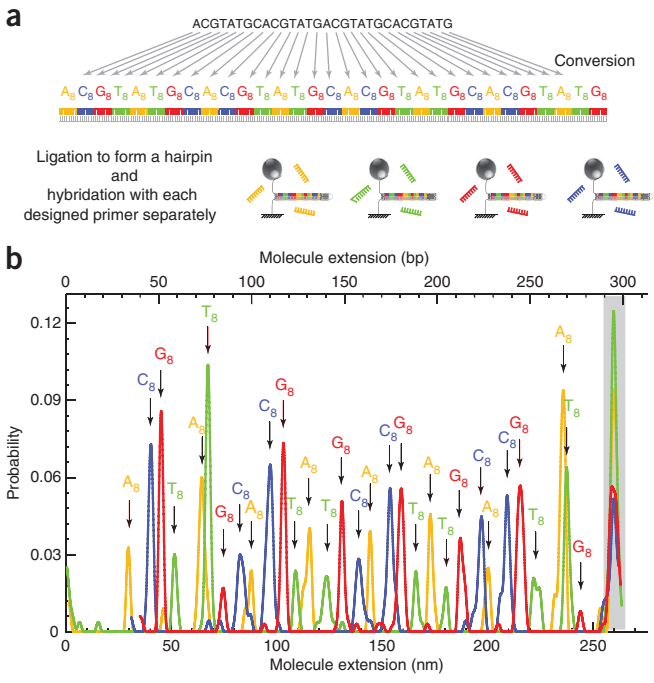

Figure 4 | Single-molecule sequencing by hybridization. (**a**) A 31-nt target DNA was converted with 8-mer A_8, T_8, G_8 and C_8 to yield a 248-bp fragment, which was then ligated to form a 295-bp hairpin (Online Methods). (**b**) To sequence the hairpin we hybridized its open state with the coding oligonucleotides. The histogram of the hairpin blockage during rehybridization at F_{test} is presented for each oligonucleotide; full extension of the molecule at F_{test} is shaded in gray.

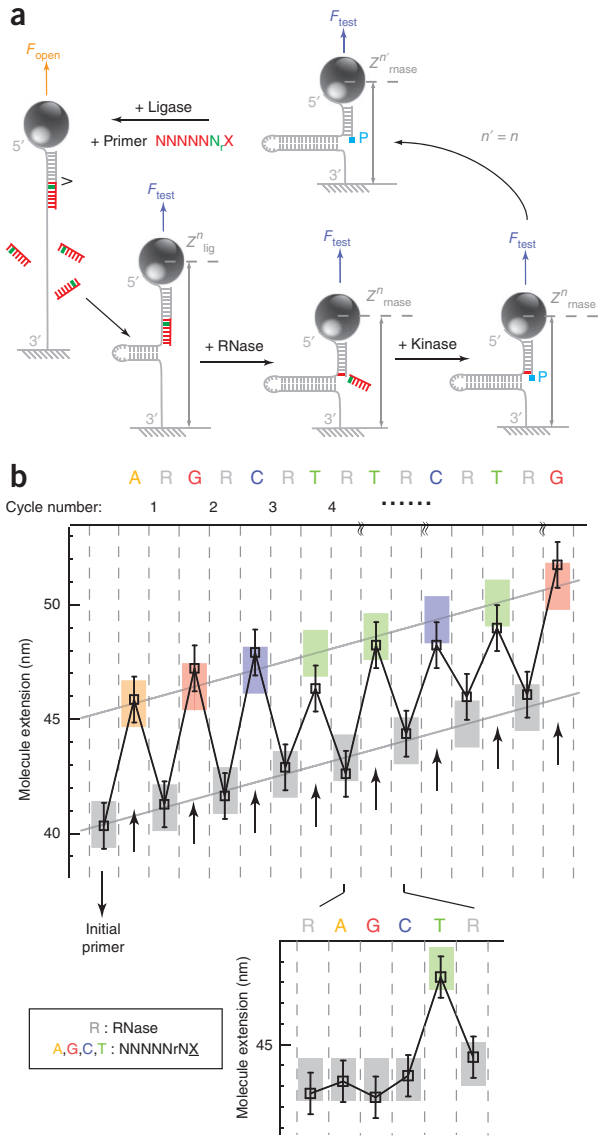

Figure 5 | Single-molecule sequencing by ligation. (**a**) Principle of sequencing by cyclic ligation and cleavage reactions. A target hairpin DNA is probed by successive cycles of injections first of oligonucleotides NNNNNN$_r$X with ligase then followed by RNase and finally by kinase. The success of the ligation is determined by the change of molecule extension before and after the ligation phase with a single ribonucleotide (green). (**b**) Sequencing of a target DNA, the expected position upon ligation is represented by colored bands: NNNNNN$_r$A (orange), NNNNNN$_r$G (red), NNNNNN$_r$C (blue) and NNNNNN$_r$T (green). The expected position after RNase cleavage is represented by gray bands. The experimental signal is plotted in black: upon ligation the molecule's extension at F_{test} increases by $Z = Z^n_{lig} - Z^{n-1}_{RNase}$ ~5 nm. Upon RNase cleavage it decreases by: $Z = Z^n_{lig} - Z^n_{RNase}$ ~4 nm. The error bar for this short hairpin is 1 nm (**Supplementary Fig. 2c**). Inset, change in extension upon injection of a mismatched oligonucleotide (ending with A, G or C; no substantial change in extension) or of a matched oligonucleotide (ending with T; extension increase of 6 nm).

primer (**Fig. 5a**). We used a 7-nt library, 5′-NNNNNN$_r$X-3′, in which X is the tested base, N represents any of the four deoxyribonucleotides and N$_r$ represents any of the four ribonucleotides. We tested the ligation to a primer strand of each of the four tested bases in hairpin opening and closing cycles. We attempted ligation in the open state of the hairpin. If successful (that is, if X is a complementary base) the oligonucleotide will block rezipping of the last seven nucleotides of the hairpin. Such an event resulted in a detectable increase in the hairpin end-to-end distance (Z^n_{lig}) of $7\alpha_{sd}$ of ~5 nm (α_{sd} being the extension change when one base pair of the hairpin is converted to ssDNA and dsDNA) (**Supplementary Fig. 2**), whereas we observed no change in the absence of ligation. This ligation step was followed by RNase cleavage of the last six nucleotides, thus finally extending the primer strand by a single base. Such cleavage allowed rezipping of the hairpin by six nucleotides with a concomitant decrease in extension (Z^n_{rnase}) with $6\alpha_{sd}$ of ~4 nm. Therefore, upon successive injections of one of the four 5′-NNNNNN$_r$X-3′ oligonucleotides, an increase of 5 nm after ligation followed by a decrease of 4 nm after cleavage was a clear indication of the incorporation of a complementary nucleotide (**Fig. 5b**).

To assess the accuracy of the ligation process, we compared the time t necessary for ligation of a 10-nt primer to an adjacent 7-nt oligomer possibly displaying a single mismatch. We did this by opening the hairpin for a time t and measuring the probability $P(t)$ of ligation (**Fig. 3**). When complementary base pairs were present, $P(t)$ saturated exponentially with a typical time τ of ~10 s: $P(t) = 1 - \exp(-t/\tau)$. With a mismatch, τ increased drastically to hours. This gives an estimation of the intrinsic ligation error (ε) when the hairpin is opened for a time t_{trial}: $\varepsilon = P_{mismatch}(t_{trial})/P_{match}(t_{trial})$. The overall error of sequencing is the sum of the intrinsic ligation error, and the error dz_m associated with the extension measurement Z: a false positive occurs if dz_m exceeds half the value of a ligation event ($Z^n_{lig}/2$ of ~2.5 nm) . With dz_m = 1 nm (**Supplementary Fig. 4**) the detection error was smaller than 0.5%. Thus the error of the presented scheme is in the range of 1% (**Table 1**).

Table 1 | Ligation error rate on template regions with 43% and 71% G+C content

Oligonucleotide	Expected: GAGCGGA	Lifetime (τ_{lig})	Error rate (t_{trial} = 40 s)
NNNGNN$_r$N	Mismatch	2,235 s	1.77%
NNNCNN$_r$N	Match	6 s	
NNNTNN$_r$N	Mismatch	8,459 s	0.47%
NNNANN$_r$N	Mismatch	27,302 s	0.15%
Oligonucleotide	**Expected: AATCAGG**	**Lifetime (τ_{lig})**	**Error rate (t_{trial} = 50 s)**
NNNGNN$_r$N	Mismatch	19,825 s	0.25%
NNNCNN$_r$N	Match	45 s	
NNNTNN$_r$N	Mismatch	3,091 s	1.60%
NNNANN$_r$N	Mismatch	22,275 s	0.22%

We tested for the ligation of a matched 10-nt oligomer to a possibly matched, adjacent 7-nt oligomer (**Fig. 3**). We fitted the measured probability of ligation $P(t)$ to an exponential form: $1 - \exp(-t/\tau)$. We thus determined the ligation time τ_{match} and $\tau_{mismatch}$ for matched and mismatched oligonucleotides. The error rate for a ligation time t_{trial} was then computed as $\varepsilon = P_{mismatch}(t_{trial})/P_{match}(t_{trial})$.

DISCUSSION

The single-molecule platform for DNA identification and sequencing presented here has several advantages over other single-molecule approaches. The hybridization fingerprinting method provides not only DNA identification but also the location of the hybridization. Single-cycle ligation allows the accurate determination of the

presence of a 16–20 nt sequence in a biological DNA sample. One can easily detect single-nucleotide polymorphisms via the ligation of a matched adjacent oligomer. Indeed, a 7-nt oligomer with the desired base mutation at its center will bind long enough to be ligated with an adjacent 10-nt oligomer and will be detected as a very long pause in the rehybridization pathway. One can directly implement several single-ligation cycles[2,3] for local resequencing without fluorescence or the need for a preamplification step. In the single-ligation scheme, the total error remains in the ~1% range, to our knowledge smaller than the error of all other single-molecule sequencing methods. A potential limitation of the method we described is the existence of secondary hairpin structures in the DNA substrate, which can be minimized (**Supplementary Discussion**).

With our instrumentation, we exchanged buffer manually, limiting our ligation sequencing method to 12 cycles and leading to the identification of only 8 consecutive bases. However, we have showed that longer read lengths are possible with an automated apparatus (**Supplementary Discussion**). We performed nine consecutive ligations of a 7-nt oligomer spanning 63 nucleotides and probing 9 bases, or 72 bases with nine 8-mers (**Supplementary Figs. 6** and **7**). Full sequencing of these 63 nucleotides would require seven similar cycles interrogating a different base shifted along the oligomer, which we have not attempted. The strand built by the successive ligations in each cycle can be removed by a denaturing agent, exonuclease or helicase (**Supplementary Fig. 7**). Furthermore, we detected the specific ligation of two target oligonucleotides on a 1,000-bp DNA (**Supplementary Fig. 4**), suggesting that even longer reads may be possible.

Using the hairpin's end-to-end distance to monitor the successive ligation of oligonucleotides has some advantages over current high-throughput sequencing schemes. These usually rely on the fluorescence detection of nucleotide addition through ligation or polymerization in a bulk sample comprised of many identical molecules. As the yield of ligation (or polymerization) is never 100%, a fraction of the DNA molecules will not incorporate the matching (oligo-)nucleotide and will become unsynchronized from the rest of the molecules; at the next cycle these molecules will present a sequence shift and incorporate a wrong (oligo-) nucleotide. This desynchronization issue, which worsens as more bases are sequenced, limits the read length of all bulk methods. In our single-molecule approach, the position of a nucleotide along a given hairpin is determined independently from previously incorporated ones; hence there are no synchronization issues. Moreover in ligation sequencing, a nucleotide incorporation is signaled by both an increase in the molecule's extension upon ligation followed by a decrease upon RNase treatment. The later step thus provides proofreading of the former.

The drawback of our method is that each nucleotide incorporation must be tested sequentially, whereas with fluorescence methods a four-color multiplexing can be used. However the fluorescence signal of a single molecule may bleach so fast that nucleotide incorporation may not be detected. This is the main cause of the large error rate of these methods. The mechanical measurement of the hairpin extension does not suffer from this problem as it can be repeated many times. Even though continuously stretching a single-molecule may break it, this appears not to be a serious issue: we routinely submit hairpins to more than 20,000 opening and closing cycles without any problems for a considerable fraction of them. When the hairpin breaks, we suspect that the weak link is between the ligand at the molecule's end (biotin or digoxigenin) and the receptor, which could be replaced by a stronger covalent bond if necessary.

The magnetic trap setup consists of a pair of rare earth magnets that provide a constant force acting simultaneously on all the beads, which is an ideal situation for a high-throughput system. Our current setup, which we implemented on a commercially available magnetic tweezers apparatus, can monitor the distance to the surface of up to 50 beads with a resolution of ~1 nm using real-time parallel image processing[26,27] (**Supplementary Fig. 8**). Recent work has enhanced the throughput to 450 beads[28]. We propose to use an evanescent illumination to further increase throughput while maintaining a low-cost charge-coupled device (CCD) or complementary metal-oxide semiconductor (CMOS) camera as a sensor (**Supplementary Discussion**). Finally, as all the molecules used here are standard oligonucleotides with no expensive fluorescent modifications, the overall cost of this platform should be minimal. In comparison to more exploratory methods, such as nanopore sequencing, for our approach in this proof-of-concept work we demonstrated real sequence identification, which has not yet been achieved with other approaches to our knowledge.

Upon automation, one should be able to use the proposed process to sequence a genome or quantify a cDNA sample. Because magnetic traps act simultaneously on many beads, our scheme is inherently parallelizable. The resolution and throughput can be readily improved by detecting the light diffracted by the beads using evanescent wave illumination[29]. Automating buffer exchanges and implementing a mechanized scanning stage are obvious improvements toward that high-throughput goal. Finally, using a ligation scheme, the change in hairpin extension is permanent, allowing for scanning and interrogating the state of a very large number of molecules.

The major advantage of our approach is in the nature of the detected signal, namely the extension of a DNA hairpin, which can be used to measure the distance between two sequences along it. This measure provides for a continuous signal, rather than a binary on or off signal as with most fluorescence-based methods. We achieved the signal with regular non–fluorescence-based imaging and standard cameras, which should provide for a low-cost alternative to current methods. Moreover, the continuous extension signal is more informative than a fluorescence signal. For instance, in the hybridization fingerprinting method, the position of the hybridization events can be detected with almost single-base accuracy. As a result, the hybridization pattern along the sequence, rather than the mere presence of hybrids, can be used as a fingerprint for a DNA sequence. This type of information is new and may open the way for applications beyond the mere adaptation of the DNA sequencing schemes demonstrated here.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturemethods/.

Note: Supplementary information is available on the Nature Methods website.

ACKNOWLEDGMENTS

We acknowledge useful suggestions by M. Volovitch, T. Lionnet and K. Neuman. This work was supported by an ERA-MolMachines grant (to D.B.), a Human Frontier Science Program grant (RGP003/2007 to V.C. and S.J.B.) and the

European Research Council grant 'MagRepS' 267862 (to V.C., F.D., M.M. and S.J.B.). We thank J. Quintas for providing mechanical expertise on the instrument.

AUTHOR CONTRIBUTIONS

J.-F.A. and V.C. designed the apparatus; F.D., S.J.B., M.M., M.M.S. and V.C. discussed the T4 system that led to the concept of sequencing; F.D. and M.M. performed experiments on magnetic tweezers; F.D. and M.M.S. prepared DNA hairpins; F.D., D.B. and V.C. performed hybridization and ligation assays; F.D. and V.C. analyzed data; F.D., M.M., M.M.S., S.J.B., J.-F.A., D.B. and V.C. prepared the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details accompany the full-text HTML version of the paper at http://www.nature.com/naturemethods/.

Published online at http://www.nature.com/naturemethods/.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Sanger, F., Nicklen, S. & Coulson, A.R. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467 (1977).
2. Drmanac, R. *et al.* Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. *Science* **327**, 78–81 (2010).
3. Shendure, J. *et al.* Accurate multiplex polony sequencing of an evolved bacterial genome. *Science* **309**, 1728–1732 (2005).
4. Braslavsky, I., Hebert, B., Kartalov, E. & Quake, S.R. Sequence information can be obtained from single DNA molecules. *Proc. Natl. Acad. Sci. USA* **100**, 3960–3964 (2003).
5. Pushkarev, D., Neff, N.F. & Quake, S.R. Single-molecule sequencing of an individual human genome. *Nat. Biotechnol.* **27**, 847–850 (2009).
6. Pihlak, A. *et al.* Rapid genome sequencing with short universal tiling probes. *Nat. Biotechnol.* **26**, 676–684 (2008).
7. Shendure, J. & Ji, H. Next-generation DNA sequencing. *Nat. Biotechnol.* **26**, 1135–1145 (2008).
8. Shendure, J., Mitra, R.D., Varma, C. & Church, G.M. Advanced sequencing technologies: methods and goals. *Nat. Rev. Genet.* **5**, 335–344 (2004).
9. Metzker, M.L. Sequencing technologies—the next generation. *Nat. Rev. Genet.* **11**, 31–46 (2010).
10. Fuller, C.W. *et al.* The challenges of sequencing by synthesis. *Nat. Biotechnol.* **27**, 1013–1023 (2009).
11. Eid, J. *et al.* Real-time DNA sequencing from single polymerase molecules. *Science* **323**, 133–138 (2009).
12. Greenleaf, W.J. & Block, S.M. Single-molecule, motion-based DNA sequencing using RNA polymerase. *Science* **313**, 801 (2006).
13. Munroe, D.J. & Harris, T.J.R. Third-generation sequencing fireworks at marco island. *Nat. Biotechnol.* **28**, 426–428 (2010).
14. Clarke, J. *et al.* Continuous base identification for single-molecule nanopore DNA sequencing. *Nat. Nanotechnol.* **4**, 265–270 (2009).
15. Treffer, R. & Deckert, V. Recent advances in single-molecule sequencing. *Curr. Opin. Biotechnol.* **21**, 4–11 (2010).
16. Husale, S., Persson, H.H.J. & Sahin, O. DNA nanomechanics allows direct digital detection of complementary DNA and microRNA targets. *Nature* **462**, 1075–1078 (2009).
17. Clark, M.D. *et al.* An oligonucleotide fingerprint normalized and expressed sequence tag characterized zebrafish cDNA library. *Genome Res.* **11**, 1594–1602 (2001).
18. Herwig, R. *et al.* Information theoretical probe selection for hybridisation experiments. *Bioinformatics* **16**, 890–898 (2000).
19. Guerasimova, A. *et al.* New tools for oligonucleotide fingerprinting. *Biotechniques* **31**, 490–495 (2001).
20. Gosse, C. & Croquette, V. Magnetic tweezers: micromanipulation and force measurement at the molecular level. *Biophys. J.* **82**, 3314–3329 (2002).
21. Brower-Toland, B.D. *et al.* Mechanical disruption of individual nucleosomes reveals a reversible multistage release of DNA. *Proc. Natl. Acad. Sci. USA* **99**, 1960–1965 (2002).
22. Strick, T.R., Allemand, J., Bensimon, D., Bensimon, A. & Croquette, V. The elasticity of a single supercoiled DNA molecule. *Science* **271**, 1835–1837 (1996).
23. Essevaz-Roulet, B., Bockelmann, U. & Heslot, F. Mechanical separation of the complementary strands of DNA. *Proc. Natl. Acad. Sci. USA* **94**, 11935–11940 (1997).
24. McNally, B. *et al.* Optical recognition of converted DNA nucleotides for single-molecule DNA sequencing using nanopore arrays. *Nano Lett.* **10**, 2237–2244 (2010).
25. Mir, K.U., Qi, H., Salata, O. & Scozzafava, G. Sequencing by cyclic ligation and cleavage (CycLiC) directly on a microarray captured template. *Nucleic Acids Res.* **37**, e5 (2009).
26. Manosas, M., Spiering, M.M., Zhuang, Z., Benkovic, S.J. & Croquette, V. Coupling DNA unwinding activity with primer synthesis in the bacteriophage T4 primosome. *Nat. Chem. Biol.* **5**, 904–912 (2009).
27. Kim, K. & Saleh, O.A. A high-resolution magnetic tweezer for single-molecule measurements. *Nucleic Acids Res.* **37**, e136 (2009).
28. De Vlaminck, I. *et al.* Highly parallel magnetic tweezers by targeted DNA tethering. *Nano Lett.* **11**, 5489–5493 (2011).
29. Singh-Zocchi, M., Dixit, S., Ivanov, V. & Zocchi, G. Single-molecule detection of DNA hybridization. *Proc. Natl. Acad. Sci. USA* **100**, 7605–7610 (2003).

ONLINE METHODS

DNA hairpin construction. Short DNA hairpins (83 bp for experiments in **Figs. 1**, **2** and **5**, 295 bp in **Fig. 4** or 179 bp in **Fig. 2** and **Supplementary Fig. 7**) were constructed by ligating three separate synthetic oligonucleotides (Eurogentec and Integrated DNA Technology; all oligo sequences are listed in **Supplementary Note**) (**Supplementary Fig. 9**). In the first step, oligos A-1 and A-2 were annealed to the complementary oligo A-3 in deionized H_2O by heating to 95 °C for 5 min, then rapidly cooling to 80 °C, followed by a slow decrease of 0.7 °C every 10 s until reaching 4 °C. These annealed products, marked part A, were cleaned up with NucleoSpin Extract II kits (Clontech). We repeated the annealing and clean up procedure for oligos B-1 and B-2, which correspond to the middle section of the 83-bp hairpin, or oligos B′-1 and B′-2 which correspond to the middle section of the 179-bp hairpin, or oligos B″-1 and B″-2, which correspond to the middle section of the 295 bp hairpin (owing to the length limit of commercial oligonucleotides, this B″ part is actually generated by the ligation of four 64-nt oligonucleotides). These annealed products were marked part B, B′ or B″. We ligated part A, part B (or part B′, B″) to oligo C, which is the loop of hairpin, using T4 ligase (5 U/µl, Fermentas) in the 1× T4 ligase reaction buffer at 25 °C for 1.5 h, then stopped the reaction by heating to 65 °C for 20 min. The ligation mixture was cleaned up with NucleoSpin Extract II kits. Finally, the digoxigenin labels were added by a fill-in reaction using Klenow(3′→5′ exo$^-$) (New England Biolabs) in the 1× NEB2 buffer with 1 mM digoxigenin-dUTP (Roche) at 37 °C for 15 min and stopped by heating to 75 °C for 20 min. The hairpin products were cleaned up with NucleoSpin Extract II kits again.

The preparation method for the 1,241-bp hairpin (**Figs. 2**, **3**, **Supplementary Figs. 2** and **6**) is described in ref. 26.

Magnetic tweezers. The magnetic tweezers device used in this work is a prototype of the commercial PicoTwist magnetic tweezers device (**Supplementary Fig. 10**). It consists of a dedicated thermally regulated inverse microscope stage coupled with a set of permanent magnets placed above the flow cell driven by translation and rotation stages. The microscope stage is designed to reduce thermal drifts; a 100×, 1.2 numerical aperture (NA) oil-immersion objective (Olympus) was focused using a piezoelectric stage with a 250 micrometer range (Pifoc PI). The objective, flow-cell holder and piezoelectric stage were all embedded in a massive metallic plate thermally regulated within 0.005 degrees.

The image of the flow cell was obtained using a 75-mm focal length field lens on a 1,320 × 1,024 pixel 31-Hz CCD camera (CM-140 GE Jai) placed below the objective.

The magnets were two Ne-Fe-B rare earth magnets of rectangular shape 2 mm × 1 mm × 1 mm glued to an iron support. The magnetic poles were horizontal in opposite directions for both magnets. The distance separating the magnets was 0.3 mm, forming a gap and generating a strong magnetic gradient.

The iron support was drilled through its center so that the light of a 665-nm LED (Luxeon Rebel) placed above the magnets passes through them illuminating the sample with a collimated beam.

The magnetic field gradient decays sharply with the distance from the magnet surface (typically a factor of 2 for every 0.3 mm). To reach high forces, the flow cell was made very thin so that the magnets can approach to 100 µm above the beads.

The flow cell was made from a glass slide (24 mm × 60 mm #1) with a 50-µm-thick piece of double-sided sticky tape with a channel cut in the middle, and the top of the flow cell was made from a 50-µm-thick mylar sheet.

The force applied to the bead was monitored by moving the magnets vertically with a translation stage (M126, PI). The software was a homemade program written in C, which tracks the *x*, *y* and *z* positions of more than 50 beads in real time. These positions together with all the experimental parameters were stored directly on the PC disk.

Forces were deduced from the distance separating the magnets from the beads, Z_{mag}. In a previous experiment, a force versus Z_{mag} function has been calibrated using long dsDNA lambda molecules and Brownian fluctuations[22] to obtain absolute calibration. For the DNA hairpin molecules, the forces deduced from this calibration were checked by measuring the force threshold of unzipping DNA. We validated this calibration process on a set of selected molecules by measuring the Brownian fluctuations using a high-speed camera that acquired video at 500 Hz (TM-6740GE). Typically force accuracy was 10–20%. The force applied to the beads varied with the magnetic field gradient and the bead magnetization. Thus for a given magnet position, the force applied on beads presents a distribution related to that of the bead magnetization. For the MyOne (Dynal, Invitrogen) beads used in this work, this parameter is relatively uniform within 20%.

Single-molecule assay. A 24 mm × 60 mm #1 glass slide was cleaned with 5 M sodium hydroxide for 2 min, then rinsed with deionized water and blown dry. A 25 µl aliquot of SigmaCote (Sigma-Aldrich) was spread on the glass and allowed to dry, then rinsed with deionized water and blown dry. While rinsing, visual inspection of the water drop contact angle attested to the surface hydrophobicity.

The flow cell was then assembled: a 50-µm-think piece of double-sided sticky tape with a 5 mm × 40 mm cut-out in the middle was adhered to the glass slide and a 50-µm-think mylar sheet to form the flow channel. Holes were punched in the mylar; a 200 µl inlet reservoir and outlet tubing were connected with plastic connectors and double-sided sticky tape. The outlet tubing was connected to a syringe pump to draw buffer through the flow cell.

Anti-digoxigenin (25 µl, 100 µg/ml, Roche) in phosphate buffered saline (PBS) was gently drawn into the chamber to avoid forming bubbles and then incubated for 2 h at room temperature. The cell was then rinsed with 500 µl passivation buffer made from PBS, bovine serum albumin (BSA) and Pluronic F127 surfactant and incubated overnight to reduce non-specific binding.

MyOne streptavidin coated beads (5 µl, 1 µm diameter, Invitrogen) were rinsed three times with 100 µl PBS using a magnetic holder to aid in separating the beads and buffer. The beads were finally resuspended in 25 µl PBS and 1 µl of DNA hairpin (typically at 50 pM), and the mixture was gently agitated. This preparation was kept on a rotating stage to prevent bead aggregation.

Bead injections were performed by using the syringe pump to produce a gentle flow of typically 10 µm/s to draw 5 µl of the bead solution placed at the flow cell inlet so that the beads quickly spread over the entire chamber. The flow was then arrested and the beads deposited on the glass surface that had been pretreated

with anti-digoxigenin. After 5–15 min, the flow cell was rinsed thoroughly with a gentle flow to remove unattached beads. The chamber was then ready for the experiment.

After these steps, magnets were juxtaposed to the beads, and a force was applied. Modulating the force between two values, 20 pN and 5 pN, distinguished beads stuck by a nonspecific interaction (which do not move) from those attached to a DNA hairpin whose extension switches between two values, Z_{open} and Z_{close}.

Once the beads were attached, the operator needed to acquire a calibration image for each bead, used to track the vertical displacement of this bead during the experiment. This calibration accurately determined the Z position of the bead in real time. The Z measurement was based on the recognition of the shape of the diffraction rings around the bead image, when it is slightly defocused. To record this calibration image, the bead was pulled vertically with a strong force to minimize fluctuations and the defocusing was scanned while the bead radial profile was recorded. During the measurements, the objective was kept at a fixed position so that the changes in the diffraction rings were caused only by the changes in extension, which were evaluated by the computer using the calibration image as a reference.

The acquisition of the molecule's extension $Z(t)$ is done with a CCD camera operating at 31 Hz. The raw data were averaged over 1 s, achieving a resolution of ~1 nm. Although the device was thermally regulated, slow drifts remained, typically on the order of 20 nm per half hour. During data collection different schemes were used to minimize these residual drifts. One scheme involves subtracting the vertical position of a reference bead stuck to the surface (Z_{ref}) from the vertical position traces $Z(t)$. Alternatively we used a differential measurement: for every opening/closing cycle, we took the center ($<Z_{close}>$) of a Gaussian fit to $Z_{close}(F_{test})$ as the reference position of the closed hairpin (extension, 0 nm). Similar Gaussian fits were used to determine the average position in one cycle of $Z_{open}(F_{test})$ and $Z_{block}(F_{test})$. The error bars represent the s.e.m.

For the hybridization identification in **Figure 2**, molecule 1 was the 1,241-bp hairpin, molecule 2 was the 83-bp hairpin, molecule 3 was the 179-bp hairpin, oligo 1 was 5′-GAAGAGACCC-3′ and oligo 2 was 5′-CAGCCGATGAC-3′. The buffer used to produce the data was the passivation buffer containing: 1× PBS, 0.2% BSA, 0.2% Pluronics surfactant, 5 mM EDTA and 10 mM sodium azide, filtered with 0.22-μm filter.

The ligation reaction in **Figure 3**, **Table 1** and **Supplementary Figures 4**, **6** and **7** was performed in ligation buffer with 0.02 U/μl T4 DNA ligase (Fermentas). For the experiment in **Figure 3**, the 10-nt primer was 5′-phos-ACAGCCAGCA-3′ and the 7-nt oligo was 5′-TAACCGA-3′. For **Table 1**, the expected AATCAGG sequence was tested with an initial primer 5′-phos-ACATACATCAG-3′ and ~2 μM oligonucleotide $NNNXNN_rN$, whereas the expected GAGCGGA sequence was tested with an initial primer 5′-phos-AACTAACCGA-3 and ~2 μM oligonucleotide $NNNXNN_rN$. In experiments shown in **Supplementary Figures 7** and **10**, the oligonucleotides ligated were 5′-phos-ACCGACCG-3′, 5′-phos-CGCTACCG-3′, 5′-phos-CACGCGCT-3′, 5′-phos-CACGCACG-3′, 5′-phos-AGCCCACG-3′, 5′-phos-AGCCAGCC-3′, 5′-phos-ACCGAGCC-3′, 5′-phos-CACGACCG-3′ and 5′-phos-ACCGCACG-3′. (Ligation buffer contained 50 mM Tris-HCl (pH 7.5), 5 mM $MgCl_2$, 1 mM DTT, 1 mM ATP and 0.2% (w/v) BSA.)

The hybridization sequencing reaction in **Figure 4** was performed in 3 M tetramethylammonium chloride (TMACL; Sigma) (ref. 30) and 0.5% (w/v) BSA with ~100 nM oligonucleotides: for the 295-bp hairpin in **Figure 4**, using A_8 (GCCACCGA), T_8 (GCACGCCA), G_8 (TCGCGCAC) and C_8 (CCGATCGC) separately. We recorded the blockage position on the complimentary hairpin stem at 25 °C.

The ligation sequencing in **Figure 5** was performed in ligation buffer for the ligation reaction. We repeatedly opened and closed the hairpin for ~10 min in the presence of ~5 μM oligonucleotide $NNNNN_rNX$ and 0.02 U/μl T4 DNA ligase (Fermentas). Then, we changed the buffer to 10 mM Tris-HCl (pH 7.5), 1 mM EDTA and 0.02 U/μl RiboShredder (Epicentre Biotechnologies) for ~10 min while reducing the stretching force to ~4 pN to increase the efficiency of the RNase. Finally, we kept the stretching force at F_{test} and flushed the previous buffer with 0.2 U/μl T4 polynucleotide kinase (Fermentas) to phosphorylate the 5′ end of the ligated and cleaved oligonucleotide. Note, RNase Blend can also work in the ligase-kinase buffer, although less efficiently. We always kept the stretching force at F_{test} when changing the buffers. All experiments were performed at 25 °C.

30. Melchior, W.B. & Von Hippel, P.H. Jr. Alteration of the relative stability of dA - dT and dG* dC base pairs in DNA. *Proc. Natl. Acad. Sci. USA* **70**, 298–302 (1973).

doi:10.1038/nmeth.1925