

# Single-molecule denaturation mapping of DNA in nanofluidic channels

Walter Reisner<sup>a,b,c,2</sup>, Niels B. Larsen<sup>b</sup>, Asli Silahatoglu<sup>d</sup>, Anders Kristensen<sup>b</sup>, Niels Tommerup<sup>d</sup>, Jonas O. Tegenfeldt<sup>c,1</sup>, and Henrik Flyvbjerg<sup>b,1</sup>

<sup>a</sup>Department of Physics, McGill University, Montreal, QC, Canada; <sup>b</sup>Department of Micro- and Nanotechnology, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark; <sup>c</sup>Department of Physics, Division of Solid State Physics, Lund University, Box 118, S-221 00, Sweden; and <sup>d</sup>Department of Cellular and Molecular Medicine, Wilhem Johannsen Centre for Functional Genome Research, University of Copenhagen, Blegdamsvej 3B, Building 24.4, Copenhagen N, Denmark

Communicated by Robert H. Austin, Princeton University, Princeton, NJ, May 22, 2010 (received for review December 27, 2009)

**Here we explore the potential power of denaturation mapping as a single-molecule technique. By partially denaturing YOYO@-1-labeled DNA in nanofluidic channels with a combination of formamide and local heating, we obtain a sequence-dependent “barcode” corresponding to a series of local dips and peaks in the intensity trace along the extended molecule. We demonstrate that this structure arises from the physics of local denaturation: statistical mechanical calculations of sequence-dependent melting probability can predict the barcode to be observed experimentally for a given sequence. Consequently, the technique is sensitive to sequence variation without requiring enzymatic labeling or a restriction step. This technique may serve as the basis for a new mapping technology ideally suited for investigating the long-range structure of entire genomes extracted from single cells.**

DNA barcoding | DNA denaturation | DNA optical mapping | nanochannel

Nanofabrication technology, developed originally for the integrated circuit industry, has over the past decade been increasingly applied to build small-scale fluidic devices, giving rise to a science of “nanofluidics.” Nanofluidics aims to exploit the ability of nanoscale devices to manipulate and analyze single molecules (1). Single-molecule analysis devices obviate the need for cloning and molecular amplification steps and will enable researchers to detect phenomena that in the past would have been obscured by ensemble averaging. In particular, nanofluidics aims to develop on-chip methods for rapidly and cheaply reading out sequence information from single large intact DNA molecules (hundreds of kbp long), with the ultimate goal of directly analyzing single genomes extracted from single cells. For example, such a single-molecule analysis capability could potentially be used to characterize the extent and dynamics of genetic heterogeneity, i.e., local variation of genotype among a population of cells from the same bacterial strain/tissue. In particular, substantial genetic heterogeneity can be present in bacterial colonies (biofilms) (2) and cancer (3), reflecting local adaptations of the cells to varying conditions within the evolving biofilm/tumor (e.g., oxygen and nutrient level). Genetic heterogeneity in these systems is believed to be a crucial factor in determining resistance to therapy.

At the heart of our device are nanochannel structures in which DNA molecules spontaneously stretch out (4–7), creating an extension of the molecule along the channel linear with molecule contour length (6). Nanochannel arrays interfaced with microfluidic loading channels serve as a highly parallel platform for performing physical mapping of DNA (5). Nano-confinement-based stretching does not require complicated setups or additional microfluidic plumbing to create an external stretching force, enabling easy automation, rapid acquisition of high statistics, and, eventually, savings for commercial realizations of the technique.

Nanochannel stretching of DNA can be combined with denaturation mapping and fluorescence microscopy to create a single-molecule optical barcoding technique. DNA can be melted, i.e., denatured from a double-stranded to a single-stranded form, by

increasing temperature or adjusting solution chemistry (8, 9). Regions of the molecule rich in A's and T's will melt at lower temperature than regions rich in G's and C's. Consequently, at certain temperatures the molecule will have ds regions interspersed with ss regions. This pattern of partial melting reveals the underlying sequence in a coarse-grained manner, limited fundamentally by the average size of the “cooperatively melting regions” (typically ~100's bp)

If the DNA is uniformly stained with a dye that unbinds when the DNA melts, local fluorescence of the melted region will decrease. Thus, the partial melting will create a grayscale barcode consisting of brighter and darker regions along the nanochannel-extended molecule. This barcode is unique to the sequence of the specific DNA molecule analyzed. The strength of this concept lies in the simplicity and robustness with which it can detect sequence variation along long (~100 kbp) DNA molecules to a resolution limited by optical diffraction to around 1 kbp. In contrast to earlier barcoding methods, which are based on restriction mapping (5), protein-nucleic acid labeling of the chain (10), or labeling through single-stranded nicks filled in with fluorescent nucleotides (4), our method requires no enzymatic pretreatments of the DNA, only an easily applied generic stain.

Our method is also distinct from currently existing DNA melting analysis techniques. Melting analysis has historically been performed by measuring the UV absorption of a DNA sample as temperature is increased (8, 9). For DNA in the 1- to 100-kbp range, the differential melting curve contains structure relating to the cooperative melting of sequence domains of varying GC content. In contrast, our technique can directly map the spatial position of these domains along an extended single DNA molecule. A second technique, denaturing gradient gel electrophoresis (DGGE), involves electrophoresis of DNA on a gel containing a gradient of a chemical denaturant. The DNA mobility will decrease during partial melting enabling separation of fragments identical in size but varying in sequence composition (11). In contrast to our technique, DGGE compares the melting temperature of different DNA fragments present in a sample: DGGE does not map spatial variation of melting along single molecules.

## Results and Discussion

The denaturation mapping experiments were performed with fused silica nanochannel devices fabricated in a clean-room facility

Author contributions: W.R., J.O.T., H.F., A.K., and N.B.L. designed research; W.R. designed the device and conducted experiments; A.S., N.T., A.K., and J.O.T. provided equipment and reagents; W.R. and H.F. analyzed data; and W.R. wrote the paper with input from J.O.T., H.F., and N.T.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

J.O.T. and H.F. were the senior authors most directly involved with helping W.R. pursue this work throughout all project stages.

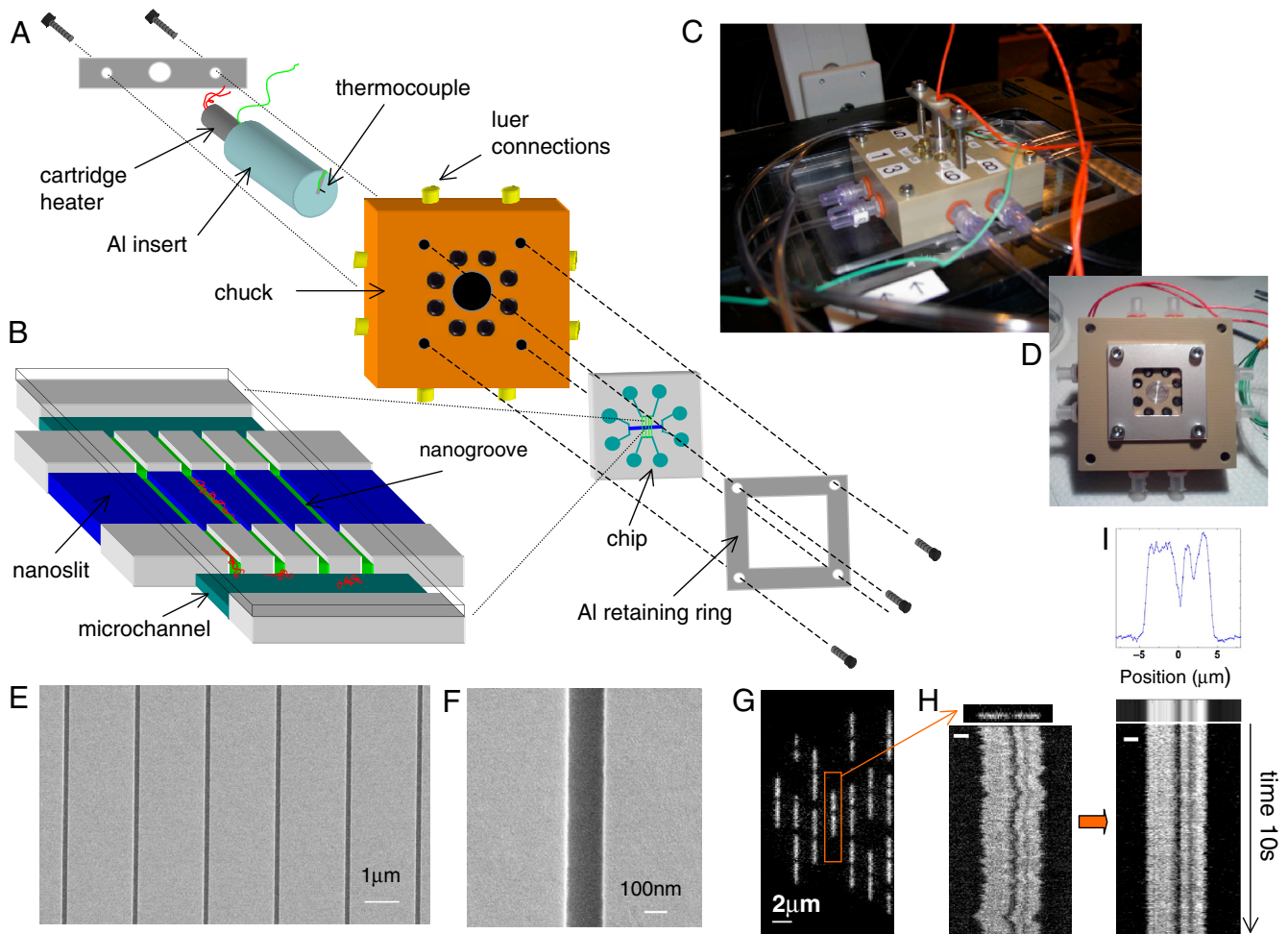
<sup>2</sup>To whom correspondence should be addressed. E-mail: reisner@physics.mcgill.ca.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1007081107/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1007081107/-DCSupplemental).

via a standard process combining electron beam and photo lithography (see *SI Text*). The devices were mounted in a chuck allowing for pressure-actuated DNA transport, heating of the device, and simultaneous optical imaging in an inverted optical microscope setup (see Fig. 1 *A–F* for experimental setup and device design). Formamide—a well-known DNA denaturant (12, 13)—was added to the running buffer to reduce the DNA melting temperature to a range accessible via optical microscopy. Formamide reduces DNA melting temperature via a simple empirical relationship:  $\Delta T_m = -0.62F^\circ\text{C}$ , where  $F$  is the percentage of formamide by volume (12). In order to minimize the heating required, we have used solutions of 50% formamide to reduce the melting temperature to just a few degrees above room temperature.

The chip is wetted in formamide-containing buffer and applied pressure is used to bring fluorescently labeled DNA molecules

from the microchannel into the nanochannel/nanogrooves (see *SI Text* for description of DNA labeling protocol and buffer chemistry). In order to increase the number of molecules available for analysis in a single field of view, we have developed a strategy for concentrating DNA in the nanogroove arrays after initial loading (see *SI Text*, Fig. S1, and Movie S1). Equally distributed pressure is applied to the four reservoirs adjoining the nanochannels to symmetrically bring DNA from the nanochannels into the nanogrooves. As the nanogrooves are deeper than the surrounding slit, at sufficiently low pressures the molecules will be trapped and accumulate in the nanogrooves. After concentration, the molecules are imaged in the nanogrooves and then ejected. This protocol (load, concentrate, image, and eject) is an efficient way to consistently maximize the number of molecules imaged at a fixed device location, lending itself to easy iteration and auto-



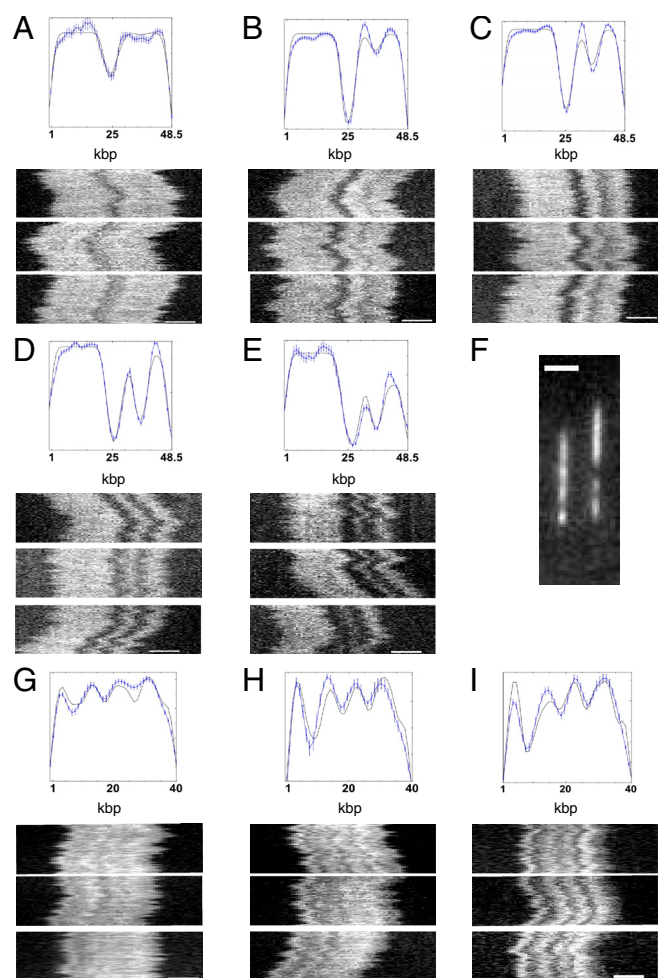
**Fig. 1.** Experimental setup and chip design. (A) Schematic of how the chip, fluidic chuck, and heater are assembled and integrated in an experiment. A retaining ring held the chip clamped against a fluidic chuck via eight o-ring seals. The chuck was designed so that pneumatic pressure could be applied to bring fresh DNA into the nanoslit from the loading channels. In addition, the chuck was designed with a built-in “hot plate” so that the chip could be heated during experiments. A cartridge heater (Omega) was inserted into a cylindrical aluminum insert that was then gently pressed against the chip’s back surface. In order to monitor the temperature in situ, a thermocouple was threaded through the insert and placed in a hole drilled halfway through the silica sample. (This hole was located 2 mm away from the nanogroove region.) The hot-plate surface was painted with thermally conductive grease (Omegatherm) to enhance heat transfer. Thermocouple measurements of the chip surface temperature show that the measured thermal difference across the chip is comparable to the expected absolute accuracy of our thermocouple readings ( $\sim 1\text{--}2^\circ\text{C}$ ). The fluctuations in thermocouple readings are  $\sim 0.1^\circ\text{C}$ . (B) Three-dimensional diagram of the nanogroove and nanochannel arrays interfaced to microscale loading channels ( $1\ \mu\text{m}$  deep,  $50\ \mu\text{m}$  wide). The nanochannels are etched first ( $120\ \text{nm}$ ), followed by the nanoslit ( $30\ \text{nm}$ ), to create nanogrooves within the open nanoslit region. Cartoon DNA molecules are shown in red. (C) Photograph of the chuck-chip-heater assembly on the microscope stage with external tubing for applying pressure. (D) Photograph of a chip clamped to the chuck via the aluminum retaining ring. (E) Low-magnification SEM micrograph of the nanochannels. (F) High-magnification SEM micrograph of a nanochannel ( $150\ \text{nm}$  wide,  $120\ \text{nm}$  deep). (G) Raw image of  $\lambda$ -DNA molecules in nanochannels at  $28^\circ\text{C}$  in 50% formamide (with background removed). (H) Example of single  $\lambda$ -DNA molecule from G with accompanying “time trace” of intensity integrated transverse to nanochannel axis plotted for all frames. (I) Time trace for molecule shown in H with thermal fluctuations “smoothed out” via rescaling procedure. (I) Time trace for molecule shown in H with thermal fluctuations “smoothed out” via rescaling procedure. The barcode shown below the plot is graphed data displayed as a grayscale plot. The scale bars are equal to  $2\ \mu\text{m}$ .

mation. Movies of nanochannel-extended molecules are obtained during the image acquisition phase of device operation (Fig. 1G). When the device temperature is elevated, the molecules exhibit characteristic banding patterns, a “barcode,” clearly visible when the intensity data are plotted as time traces (Fig. 1H).

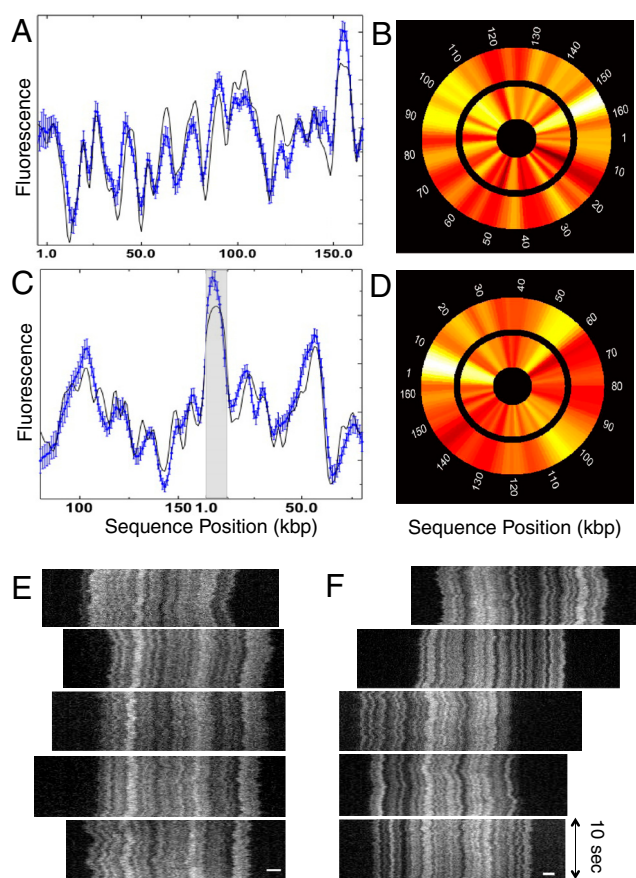
Thermal fluctuations in contour density and center-of-mass diffusion create visible distortions of the raw barcodes. In order to remove these distortions, we use an analysis procedure that aligns the center of mass in time and “smoothes out” internal density fluctuations (Fig. 1I, *SI Text*, and Fig. S2 A and B). The corrected traces are then averaged to create a single-molecule intensity profile (Fig. 1I). In order to compare and average the intensity profiles for an ensemble of molecules, we use an algorithm to systematically align the individual profiles (*SI Text* and Fig. S2 C–E). The aligned profiles are then averaged to create “consensus barcode profiles.”

As a proof of principle demonstration of our technique we have applied our methodology to four DNA sequences,  $\lambda$ -phage, T7, T4GT7, and a BAC construct from chromosome 12 (RP11-

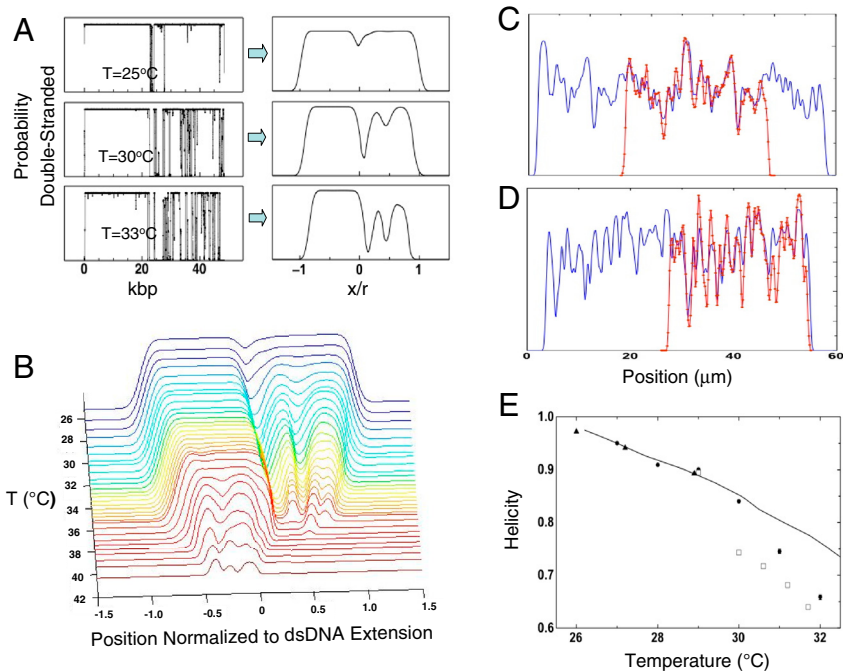
125C7) with an 11.6-kbp cloning vector. The  $\lambda$ -phage and T7 sequences (Fig. 2) give rise to distinct barcode profiles with characteristic temperature dependence. The  $\lambda$ -phage DNA develops a single central dip that grows as the temperature is increased. The T7-phage sequence develops a three-dip structure. The profiles are distinct and can easily be distinguished, even for single molecules side by side in the nanochannels (Fig. 2F). The T4GT7 profiles can be aligned, but it is clear that they must be shifted relative to each other in order to achieve alignment: The barcode appears to be circularly permuted between molecules (Fig. 3F). This agrees with the classic result of phage genetics that T4 has a linear, but circularly permuted, genome (14). The circularly permuted T4GT7 sequence is an excellent test of our alignment procedure: We can align the permuted T4GT7 barcodes to construct a circular consensus melting map (Fig. 3A and B). In a second test of our alignment procedure, the RP11-125C7 construct was run in the devices *prior* to linearization of its plasmid structure. While circular DNA cannot be easily mapped, a certain fraction of the circular BACs will break



**Fig. 2.**  $\lambda$ -phage and T7 consensus barcodes in 50% formamide. (A)  $\lambda$ -phage barcode at 27 °C with three example time traces ( $n = 13$  molecules total in consensus). (B)  $\lambda$ -phage at 28 °C ( $n = 56$  molecules in consensus). (C)  $\lambda$ -phage at 30 °C ( $n = 47$  molecules in consensus). (D)  $\lambda$ -phage at 31 °C ( $n = 38$  molecules in consensus). (E)  $\lambda$ -phage at 32 °C ( $n = 19$  molecules in consensus). (F) A raw fluorescence micrograph showing  $\lambda$ -phage and T7 molecules side by side: It is possible to easily distinguish the two from the melting barcode. (G) T7 at 30 °C ( $n = 20$  molecules in consensus). (H) T7 at 30.6 °C ( $n = 10$  molecules in consensus). (I) T7 at 31.2 °C ( $n = 17$  molecules in consensus). The example time traces have 10-s duration. The scale bar in all images is 2  $\mu$ m. The error bars in the plots are standard errors on the ensemble mean (SEM), calculated as the SD of ensemble values divided by  $\sqrt{n}$ .



**Fig. 3.** Consensus barcodes for T4GT7 and BAC RP11-125C7. (A) Consensus barcode for T4GT7 DNA (blue) with theoretical fit (bold curve). Data acquired with 50% formamide at 27 °C. Consensus obtained from 26 molecules. (B) T4GT7 barcode displayed as a circular wheel (color corresponds to grayscale value, white maximum values, dark-red minimum). The inner wheel is the theoretical prediction; the outer wheel is the experimental consensus barcode. (C) Consensus barcode for BAC RP11-125C7 (blue) with theoretical fit (bold curve). Data acquired with 48% formamide at 27 °C. Consensus obtained from 32 molecules. The shaded region is the location of the cloning vector. (D) BAC RP11-125C7 barcode displayed as a circular wheel. The inner wheel is the theoretical prediction; the outer wheel is the experimental consensus barcode. The error bars in the plots are standard deviations of the mean arising from profile variation across the ensemble. (E) Raw time traces for RP11-125C7 arranged to emphasize barcode overlap. The scale bar is 2  $\mu$ m. (F) Raw time traces for T4GT7 arranged to emphasize barcode overlap. The scale bar is 2  $\mu$ m.



**Fig. 4.** Theoretical prediction of barcode structure from dsDNA melting probability (A). Theoretical melting traces can be obtained directly from melting-probability calculations by assuming values for ds/ss DNA stretching, fluorescence, and convolving the profile with a Gaussian to simulate optical broadening. (B) Evolution of  $\lambda$ -phage melting traces with temperature across the melting transition: by scanning temperature, both high AT and high GC regions are imaged. Details in high AT regions are visible at low temperature, whereas at higher temperature details in high GC regions can be imaged. (C) Experimental barcode profile (red) for single BAC RP11-125C7 molecule aligned to theoretical estimate of RP11-125C7 barcode (blue). (D) Experimental barcode profile (red) for single T4GT7 molecule (red) aligned to theoretical estimate of T4GT7 barcode (blue). The theoretical estimates for T4GT7 and BAC RP11-125C7 are created from two identical and adjoining sequences to capture trace permutation. (E) Fitted  $\lambda$ -phage helicities vs. temperature; circles, open squares, and triangles indicate measurements with different chips/setups. The variation between experimental runs arises primarily from variations in thermocouple calibration ( $\sim 1^\circ\text{C}$  error). The bold curve is a theoretical estimate of the helicity at 5 mM NaCl.

during chip loading. These linearized BACs were then aligned and converted into a circular map (Fig. 3 C and D). In addition, the ability to align the barcodes of different molecules allows for relaxation of the requirement of DNA monodispersity. DNA fragments can be aligned to positions on overlapping molecules via the barcodes without a priori knowledge of their size and sequence position (Fig. S2 C–E).

The observed barcode structure in the DNA fluorescence profile is consistent with the presence of local denaturation. Computer algorithms exist, based on the Poland–Scheraga model (8, 9), which can predict the sequence-dependent melting probability of dsDNA  $p_{\text{ds}}(s)$  for a particular temperature and salt concentration (15). These codes use as inputs the measured enthalpies and entropies of the 10 independent base-pair duplexes (16), resulting from a combination of hydrogen bonding between bases and stacking interactions between nearest-neighbor base pairs. In addition, the codes contain a description of polymer looping energetics to correctly distinguish between internally melted regions (bubbles) and melted end segments (17).

These calculated melting probabilities can be transformed into an expected fluorescence profile (Fig. 4A). We argue that an approximate average sequence-dependent intensity  $I(s)$  and extension  $r(s)$  along the channel can be obtained from

$$I(s) = I_{\text{ds}}p_{\text{ds}}(s) + I_{\text{ss}} * [1 - p_{\text{ds}}(s)], \quad [1]$$

$$r(s) = r_{\text{ds}}p_{\text{ds}}(s) + r_{\text{ss}} * [1 - p_{\text{ds}}(s)]. \quad [2]$$

The quantity  $I_{\text{ds}}$  is the intensity of the double-stranded regions,  $I_{\text{ss}}$  the intensity of single-stranded regions,  $r_{\text{ds}}$  is the average extension per base pair of dsDNA along the channel, and  $r_{\text{ss}}$  is the average extension per base pair of ssDNA. The intensity profile  $I(r)$  obtained from this calculation is our theory for the emitted intensity. The observed intensity is the emitted intensity convoluted with the point-spread function of the microscope objective. We approximate the latter with a Gaussian with standard deviation  $\sigma \sim 200$  nm. The results of this calculation for the example case of  $\lambda$ -DNA, using the algorithm developed by Hovig and coworkers (15), are shown in Fig. 4A. The melting-probability calculation uses as inputs the 10 independent enthalpies and entropies of the stacked and paired base-pair duplexes measured by Blake and

Delcourt (16) with the loop entropy weight suggested by Blossey and Carlon (17). The computed profiles clearly reproduce qualitatively the structure and temperature dependence evident in the  $\lambda$ -DNA consensus profiles. The theoretical profiles can be directly fitted to the experimental data upon appropriate specification of  $I_{\text{ds}}$ ,  $I_{\text{ss}}$ ,  $r_{\text{ds}}$ ,  $r_{\text{ss}}$ , and  $\sigma$ . The results of the fitting procedure for  $\lambda$ -DNA and T7 DNA are shown in Fig. 2.

Using the barcode alignment technique, circularly permuted sequences can also be aligned to an estimated theoretical barcode—see Fig. 4 C and D—enabling us to compare the T4GT7 and BAC construct data to theory and calibrate the  $x$  axis in units of kbp referenced to a known position on the true sequence (see *SI Text*). The theoretical profiles convincingly capture the structure of the experimental barcodes, suggesting that the alignment procedure is effective. For example, we are able to map the position of the 11.6-kbp vector sequence, which results in the strong central peak, due to its relatively higher GC content (47.4%, compared to 37% on RP11-125C7).

The key fitting parameters are the ratio of the ssDNA to dsDNA extension,  $r_{\text{ss}}/r_{\text{ds}}$ , and the ratio of the intensities,  $I_{\text{ss}}/I_{\text{ds}}$ . We expect that  $r_{\text{ss}}/r_{\text{ds}} < 1$  because ssDNA is more flexible than dsDNA. At the ionic strength of 5 mM used in the experiment, the persistence length of dsDNA  $P_{\text{ds}}$  lies in the range of 60–80 nm (18, 19) and the persistence length of ssDNA  $P_{\text{ss}} = 6 \pm 2$  nm (20). The extension of DNA in a nanochannel is  $r \sim P^{1/3}$ , so a crude estimate is that  $r_{\text{ss}}/r_{\text{ds}} \sim (P_{\text{ss}}/P_{\text{ds}})^{1/3} \sim 0.4$ –0.5. In fact, we have found in solutions of 50% formamide that the best fits are obtained with  $r_{\text{ss}}/r_{\text{ds}} \sim 0.7$ –0.9. This result is plausible as formamide is known to extend ssDNA, a fact well established from early electron microscopy studies (21). For definiteness, in the theoretical calculations we have fixed  $r_{\text{ss}}/r_{\text{ds}} = 0.85$ , the average value obtained from the  $\lambda$ -phage fits.

Least-squares fitting yields values of  $I_{\text{ss}}$  consistent with zero. Consequently, we have fixed  $I_{\text{ss}} = 0$  for all the plots shown. Physically, this implies that the stain completely leaves the melted regions of the molecule. While it is known that the ratio of the ssDNA and dsDNA quantum yields *in bulk* is 0.66 (22), it is likely that electrostatic dye-silica surface interactions might promote the loss of stain from single-stranded regions of the molecule (we have, in fact, observed stain binding to the nanochannel surface during the melting cycle). From an application point of view, this implies that the contrast between the melted and



8. Wartell M, Benight A (1985) Thermal denaturation of DNA molecules: A comparison of theory with experiment. *Phys Rep* 126:67–107.
9. Gotoh O (1983) Prediction of melting profiles and local helix stability for sequenced DNA. *Adv Biophys* 16:1–52.
10. Chan YC, et al. (2004) DNA mapping using microfluidic stretching and single-molecule detection of fluorescent site-specific tags. *Proc Natl Acad Sci USA* 14:1137–1146.
11. Fischer SG, Lerman LS (1983) DNA fragments differing by single base-pair substitutions are separated in denaturing gradient gels: Correspondence with melting theory. *Proc Natl Acad Sci USA* 80:1579–1583.
12. Sadhu C, Dutta S, Gopinathan KP (1984) Influence of formamide on the thermal stability of DNA. *J Biosci* 6:817–821.
13. Blake RD, Delcourt SG (1996) Thermodynamic effects of formamide on DNA stability. *Nucl Acid Res* 24:2095–2103.
14. Hershey AD (1969) Idiosyncrasies of DNA structure. *Nobel lecture*.
15. Tøstesen E, Jerstad GI, Hovig E (2005) Stitchprofiles.uio.no: Analysis of partly melted DNA conformations using stitch profiles. *Nucl Acid Res* 33:W573–W576.
16. Blake RD, Delcourt SG (1998) Thermal stability of DNA. *Nucl Acid Res* 26:3323–3332.
17. Blossey R, Carlon E (2003) Reparametrizing the loop entropy weights: Effect on DNA melting curves. *Phys Rev E* 68:061911.
18. Baumann CG, Smith SB, Bloomfield VA, Bustamante C (1997) Ionic effects on the elasticity of single DNA molecules. *Proc Natl Acad Sci USA* 94:6185–6190.
19. Wenner JR, Williams MC, Rouzina I, Bloomfield VA (2002) Salt dependence of the elasticity and overstretching transition of single DNA molecules. *Biophys J* 82:3160–3169.
20. Tinland B, Pluen A, Sturm J, Gilbert W (1997) Persistence length of single-stranded DNA. *Macromolecules* 30:5763–5765.
21. Banfield HY, Inman RB (1974) The electronmicroscopy of DNA. *Annu Rev Biochem* 43:605–619.
22. Cosa G, Focsaneanu KS, McLean JRN, McNamee JP, Scalano JC (2001) Photophysical properties of fluorescent DNA-dyes bound to single- and double-stranded DNA in aqueous buffered solution. *Photochem Photobiol* 73:585–599.
23. Fang L, et al. (2007) Human genome melting map. *PLOS Comput Biol* 3:0874–0886.
24. Lin J, et al. (1999) Whole-genome shotgun optical mapping of *Deinococcus radiodurans*. *Science* 285:1558–1562.
25. Cao H, Tegenfeldt JO, Austin RH, Chou SY (2002) Gradient nanostructures for interfacing microfluidics and nanofluidics. *Appl Phys Lett* 81:3058–3060.
26. Morton KJ, Loutharback K, Inglis DW, Tsui OK, Sturm JC (2008) Crossing microfluidic streamlines to lyse, label and wash cells. *Lab Chip* 8:1448–1453.