Automation of a single-DNA molecule stretching device

Sørensen, Kristian Tølbøl; Lopacinska, Joanna M.; Tommerup, Niels; Silahtaroglu, Asli; Kristensen, Anders; Marie, Rodolphe

Published in:
Review of Scientific Instruments

Link to article, DOI:
10.1063/1.4922068

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Automation of a single-DNA molecule stretching device
Kristian Tølbøl Sørensen, Joanna M. Lopacinska, Niels Tommerup, Asli Silahtaroglu, Anders Kristensen, and Rodolphe Marie

Citation: Review of Scientific Instruments 86, 063702 (2015); doi: 10.1063/1.4922068
View online: http://dx.doi.org/10.1063/1.4922068
View Table of Contents: http://scitation.aip.org/content/aip/journal/rsi/86/6?ver=pdfcov
Published by the AIP Publishing

Articles you may be interested in
An automated microfluidic system for single-stranded DNA preparation and magnetic bead-based microarray analysis
Biomicrofluidics 9, 024102 (2015); 10.1063/1.4914024

High throughput fabrication of disposable nanofluidic lab-on-chip devices for single molecule studies
Biomicrofluidics 6, 036502 (2012); 10.1063/1.4740231

An improved microfluidics approach for monitoring real-time interaction profiles of ultrafast molecular recognition

A robotics platform for automated batch fabrication of high density, microfluidics-based DNA microarrays, with applications to single cell, multiplex assays of secreted proteins
Rev. Sci. Instrum. 82, 094301 (2011); 10.1063/1.3636077

Plasma thinned nanopost arrays for DNA electrophoresis
J. Vac. Sci. Technol. A 29, 011025 (2011); 10.1116/1.3531693
Automation of a single-DNA molecule stretching device

Kristian Tølbøl Sørensen, Joanna M. Lopacinska, Niels Tommerup, Asli Silahtaroglu, Anders Kristensen, and Rodolphe Marie

1Department of Micro- and Nanotechnology, Technical University of Denmark, Kongens Lyngby, Denmark
2Department of Cellular and Molecular Medicine, University of Copenhagen, København, Denmark

(Received 27 March 2015; accepted 21 May 2015; published online 9 June 2015)

We automate the manipulation of genomic-length DNA in a nanofluidic device based on real-time analysis of fluorescence images. In our protocol, individual molecules are picked from a microchannel and stretched with pN forces using pressure driven flows. The millimeter-long DNA fragments free flowing in micro- and nanofluidics emit low fluorescence and change shape, thus challenging the image analysis for machine vision. We demonstrate a set of image processing steps that increase the intrinsically low signal-to-noise ratio associated with single-molecule fluorescence microscopy. Furthermore, we demonstrate how to estimate the length of molecules by continuous real-time image stitching and how to increase the effective resolution of a pressure controller by pulse width modulation. The sequence of image-processing steps addresses the challenges of genomic-length DNA visualization; however, they should also be general to other applications of fluorescence-based microfluidics. © 2015 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution 3.0 Unported License. [http://dx.doi.org/10.1063/1.4922068]

I. INTRODUCTION

Nucleic acid analysis in lab-on-a-chip devices has been the focus of intense research. Microfluidics has emerged as a key technology that enables extraction, amplification, and detection of nucleic acids in lab-on-a-chip format, which is relevant for a range of applications such as pathogen detection or whole genome analysis. In particular, the DNA sequence analysis via fluorescence imaging of stretched genomic-length DNA has emerged as a method complementary to DNA sequencing with the perspective of analysing single DNA molecules.

Genomic-length DNA was previously imaged on surfaces or by tethering the molecules end to a bead and stretching it using a force applied by AFM, a magnetic field, optical tweezers, or the drag of a liquid flow. More recently, stretching DNA molecules without tethering the molecule ends was demonstrated in nanofluidic devices with several possible device architectures: nanochannel devices utilising confinement alone, devices using entropic recoil as an additional force, or devices relying on the hydrodynamic drag of an elongation flow. Nanochannel devices for sizing genomic-length DNA have been automated using fluorescence detection. Chou et al. sized single DNA molecules by measuring the width of the fluorescence pulse induced by individual molecules passing through an excitation laser spot. A similar technique was later used to sort single DNA molecules based on the detection of fluorescent epigenetic markers. More recently, Hastie et al. automated a nanochannel device for stretching DNA by confinement and reported a throughput of 5 Gb/hr. The device includes imaging capabilities for imaging fluorescence barcodes, although it might not have been used for the feedback of the automation.

We recently demonstrated a device combining flow stretching and confinement using a cross-shaped shallow channel called a nanoslit. In our device, a fragment of genomic-length DNA is elongated in a symmetrical double elongation flow within the cross-shaped nanoslit. We have demonstrated that the hydrodynamic drag applies a force on the DNA in the pN range that is sufficient to stretch DNA to within a few percent of its full contour length. The high tension of the DNA strand as well as the two-dimensional confinement of the nanoslit represses the natural Brownian motion of the DNA strand during imaging. This enables reading out a sequence-specific fluorescence barcode with an unprecedented resolution, allowing structural information to be obtained from individual molecules. In addition, the device architecture allows the imaging of DNA molecules several megabase (Mb) pairs in length. The device is easily fabricated in silica or in polymer without sub-micrometer lithography process as opposed to the fabrication of 45 × 45 nm silica channels requiring e-beam lithography and material regrowth. In our device, however, high stretching of long molecules is achieved at the cost of a device architecture that limits parallelization and currently requires extensive training for manual operation.

In this paper, we describe the key automation steps of the operation of the flow-stretch architecture in an effort to make this device simpler to use and develop the optical mapping of Mb-long DNA as a tool to assist sequencing. The LabVIEW-algorithms developed for this purpose are described and demonstrated as proofs of principle. Specifically, we solve the challenges of automatically (I) selecting a single long DNA strand in a microchannel among many short DNA fragments, (II) isolating a selected DNA strand in a nanoslit, and (III) elongating an isolated DNA molecule spanning the whole length of the nanoslit. These three challenges will subsequently be referred to as phase I, II, and III, as they constitute the three main sequential phases of the experimental protocol. The demonstrated algorithms are all incorporated...
into a multi-loop application framework, such that all three phases can be performed by a single virtual instrument (VI). Many of the solutions presented here, which were tailored to the specific challenges of manipulating single DNA strands, will be general to many other fluorescence-based automation challenges.

II. METHOD

A. Experimental setup

The nanofluidic device for identifying, isolating, and elongating single DNA molecules was previously reported by Marie et al. The device design was slightly modified to allow fabrication by injection moulding in cyclic olefin copolymer (COC, TOPAS Advanced Polymers GmbH), for an all polymer version of the device. Briefly, the microchannels for sample transport were 50 µm wide and 5 µm deep pairwise connecting eight inlets, whereas the cross-shaped nanoslit was 20 µm wide and ~110 nm deep, connecting the four microchannels as shown in Figures 1(a) and 1(b). The microchannels had a constriction by the nanoslit entrances, pre-stretching DNA via an elongation flow in front of the entrance to the nano slit. The device was sealed with a 150 µm polymer lid that allows high-NA imaging using oil immersion objectives.

The nanofluidic device was mounted on the motorised stage (Prior Scientific Proscan III) of an inverted microscope (Nikon TS2000), equipped with an oil immersion objective (Nikon Plan Apo VC 60x/1.40), a 1.5x optivar lens, and an electron multiplying charged couple device (EMCCD) camera (Photometrics Evolve 512 with PVCam 2.7 driver) as depicted in Figure 1(c). The connected shutter and focus wheel were also motorised via the ProScan III-controller. The device was photo-bleached for up to 90 min, in order to reduce auto-fluorescence of the COC polymer. The inlets of the device were loaded with buffer and YOYO-1-labelled metaphase chromosomes, and megabase-sized DNA fragments were released on-chip by proteolysis, as previously reported. The inlets of the device were connected to an air pressure control instrument (Fluiagent MFCS, 8 channels, 0-1 bar range, 1 mbar resolution). The automation was implemented in a single LabVIEW 2011 (National Instruments) VI on a PC (Intel Core2Duo at 3.00 GHz and 4 GB RAM), running Microsoft Windows 7. The VI controlled the XY-stage, shutter, focus wheel, EMCCD camera, a custom-made temperature controller, and the pressure controller.

B. Device operation

The experimental protocol and the challenges that make it demanding to both manual operator and automation are described below.

First, in the “selection” phase, DNA is transported through a microchannel using pressure-driven flow. The flowing DNA is imaged live in fluorescence, and the video feed is constantly monitored (whether by operator or algorithm), until a strand of suitable size is identified among the many short DNA fragments (Figure 2(a)). The strand should at least be long enough to be able to span the 440 µm long nanoslit once fully elongated, but the precise length of the strand is otherwise unimportant. In the microchannel, where the strand is only partially stretched, this minimum length typically corresponds to 2-3

FIG. 1. (a) The device comprises a cross-shaped nano slit connecting four microchannels. (b) Microchannels and nano slit are 5 µm-deep and 110 nm-deep, respectively. (c) Schematic illustrations of the experimental setup comprising the polymer device mounted on a x-y-z stage, the high NA objective, the EMCCD camera, the stage, and pressure controller. Automation is implemented in LabVIEW. The illumination (not drawn) is comprised in the epifluorescence setup of the inverted microscope.

FIG. 2. Automated steps of single-molecule stretching. (a) A DNA molecule 3 times longer than the field of view (green box) is identified in the microchannel. (b) The molecule is introduced from (1) microchannel into (2) nanoslit, via PID-stabilised pressure ramp-up. (c) The isolated DNA molecule is stabilised in real-time and elongated using a perpendicular flow. DNA is shown schematically in red, where blue arrows indicate pressure direction and magnitude. (d)-(f) Sample frames corresponding to (a)-(c), where (f) is made by stitching 7 images.
times the field of view width of 90 μm. This phase is challenging because

- the signal-to-noise ratio (SNR) of single DNA molecules is rather low,
- molecules may only be partially in focus due to the narrow focus depth of the high-N.A. objective,
- molecules may be moving in and out of focus,
- long suitable molecules are surrounded by short unsuitable fragments, and
- flow velocity across the channel varies due to the parabolic flow profile.

The “isolation” phase follows once a DNA strand has been selected, which is then stabilised near the nanoslit entrance (Figure 2(b) and supplementary figure S1). To force the molecule into the nanoslit, the pressure must be increased by several hundred millibars, which is challenging because

- the pressure becomes unstable every time the pressure setpoint is modified, and
- even slight pressure differences can cause the molecule to escape the field of view.

Finally, in the “stretching” phase, the inserted DNA molecule is elongated via a perpendicular elongation flow from the side-arms of the cross-shaped nanoslit (top and bottom in Figure 2(c), corrective flow exemplified on the right-hand side). This phase is challenging because

- the molecule is constantly changing shape and
- elongation must be symmetrical to avoid pushing the strand out of the nanoslit, while at the same time,
- molecules are too long to fit in the field of view, so the stage must be translated, but
- the background intensity varies at different positions.
- Also,
- photo-nicking causes strand fragmentation after some time, unless the elongation is performed swiftly.

Once full elongation is achieved, images are acquired along the full length of the nanoslit, and the DNA molecule may then be flushed out of the nanoslit to a waste outlet or to a collection outlet for further analysis. The protocol then repeats from the first phase.

### III. RESULTS AND DISCUSSION

#### A. Machine vision on single molecules

Machine vision is at the core of the automated system. In every phase of the experiment, the system must locate the DNA molecule of interest and make decisions based on machine vision.

Even with a state-of-the-art EMCCD camera, fluorescence intensity from a single DNA molecule is at a premium and depends strongly on degree of staining, stretching, and focus. The degree of staining is limited by saturation of DNA, and some fluorophores must be released during denaturation, in order for the optical readout to carry information. Moreover, increasing the staining ratio or the excitation light intensity increases the rate of photo-nicking, which must critically be avoided in our application. In other words, the SNR of YOYO-1-labelled DNA molecules is limited.

We have developed a sequence of image processing steps which facilitate the detection of single-molecule DNA in a low-SNR environment and which may be helpful in many other applications of epifluorescence microscopy, where SNR is a limiting factor. Our method enhances fluorescence images and reduces image complexity into a 1-bit binary representation suitable for machine vision, by exploiting the fact that the signal from genomic DNA is highly anisotropic, while the background is isotropic.

Each raw 16-bit grayscale micrograph (Figure 3(a)) contains only vague hints of a DNA molecule, but is enhanced by applying a number of image processing steps (Figures 3(b)-3(g)), namely, a 5 frame averaging, 2 × 1 binning, background subtraction, 5 × 5-neighbourhood median filtering, anisotropic (20 × 1) median filtering, and finally binarisation.

Frame averaging is only used in the nanoslit, where motion of the DNA is slower due to the high hydraulic resistance. Movement is much more rapid in the microchannel, where fluorescence intensity on the other hand is increased by DNA coiling. Binning pools together photo-electrons from adjacent pixels on the camera CCD, increasing sensitivity at the cost of resolution. It should be noted that binning is performed on-camera prior to electron multiplication, thus increasing photo-sensitivity more than software-binning in post-processing could. Using 2 × 1 binning halves the resolution along the channel, thus emphasising oblong objects and silencing smaller fragments. The reduced resolution along the channel is an advantage in cases where a molecule is partly out of focus and could otherwise appear discontinuous. The anisotropic binning maintains full resolution across the channel in order to separate parallel DNA strands. Median filtering reduces shot-noise and the anisotropic filtering further emphasises DNA. Using the resulting binary image as a mask for the other images, the mean (μbg) of the signal and standard deviation (σbg) of the background are calculated for each processing step A-F and used to calculate SNRs in Figure 3, using the definition $SNR = 20 \cdot \log_{10} \left( \frac{\mu_{sig}}{\sigma_{bg}} \right)$.

The SNR increases in each step from B through F, in this case except for step D, as background subtraction is only advantageous when the background is inhomogeneous. The
demonstrated processing pipeline generates a vision result every ~100 ms, forming the real-time feedback upon which all downstream automation depends.

B. Phase 1: Selection

During the first phase of the experiment, a DNA fragment must be identified, which is long enough to span the nanoslit once fully stretched. DNA is partially sheared in the parabolic flow profile inside the microchannel, so fragments with an apparent length of ~220 µm are typically able to span the 440 µm long nanoslit. However, the field of view is only 91 µm wide.

The manual operator intuitively estimates the length of a DNA strand by judging how fast it moves and for how long it is visible within the field of view. The human brain excels at pattern recognition in a low-SNR environment, so keeping track of a DNA strand while many short fragments flow by is possible to the trained operator.

A similar task was solved by Chou et al.\(^\text{15}\) by measuring the fluorescence burst duration of genomic DNA stretched in a 500 × 250 nm nanochannel, by passing through a laser spot monitored by a photodetector, and by estimating molecule length based on signal duration. This approach is not suitable for our microchannels (50 × 5 µm), as DNA strands of different lengths flow in parallel.

Our solution is to construct a continuous 1-bit “flow panorama,” which can be arbitrarily long, limited only by computer memory. This is done by end-fusing a slice from each incoming video frame (Figure 4), where the slice width depends on the flow velocity, which must be determined in real-time. Standard methods for determining flow velocity, such as optical flow algorithms,\(^\text{23}\) were found to perform poorly in this setup due to the low SNR. Instead, we developed a particle tracking algorithm, which correlates objects between frames based on size and relative position, in order to measure an average displacement of the entire liquid column.

Our system did not correct for the parabolic flow velocity profile, because the contour length only had to exceed a certain threshold, and so its precise value was not a primary concern. However, this could easily be implemented using the equation for the velocity vector field \(\mathbf{v}(y, z)\) of Poiseuille flow in a rectangular channel. In the particular setup for high-NA fluorescence imaging, the depth of field is merely \(z_d \approx 320 \text{ nm}\). This is small compared to the channel height of \(h = 5 \mu\text{m}\); thus, the flow velocity within a field of view may be considered as a function of the distance to the channel wall only.

The method reported here enables the measurement of object length, without both end-points ever being visible in a field of view at the same time. As an example, the flow-panorama in Figure 4(h) shows a DNA molecule of estimated length 286 µm, more than three times the maximum length visible within the field of view. The method has proven effective at automatically identifying DNA strands of length sufficient for spanning the entire nanoslit once elongated.

C. Phase 2: Insertion

In the second phase, a selected DNA molecule is forced into the device nanoslit. This is achieved by increasing the pressure in the microchannel, while keeping the megabase-long molecule of interest in front of the nanoslit entrance. When done manually, the trained operator would increase pressure in relatively large increments (tens of millibars), then spend some time stabilising the molecule at the nanoslit before incrementing again. Despite the chosen molecule spanning multiple fields of view in its pre-stretched form (Figure 2(d)), during pressure increase, the DNA tends to coil up into a blob (Figure 2(e)), which typically exits the field of view rapidly at high pressure difference. Thus, it is impractical to increase pressure in larger increments, as this induces a large transient pressure difference between the microchannel inlets. A decimal-point pressure value would be required to stabilise the molecule exactly at the entrance, but as the pressure controller has a 1 mbar pressure resolution, and 1 mbar pressure difference usually causes the strand to change direction of movement, the operator must rapidly alternate between increasing/decreasing the pressures on one channel in order to maintain the strand position near the entrance. After some time (typically a few minutes, depending on the initial distance to the entrance), the DNA is forced into the nanoslit.

To automate this, a custom proportional-integral-derivative (PID) controller was developed and integrated into the VI, which returns a decimal pressure value based on the spatial offset between the current positions of the molecule compared to the nanoslit entrance. The PID feedback loop enables stabilisation of the selected DNA strand while the system gradually ramps up the pressure on both channels. With proper gain tuning, the PID feedback loop enabled an
object of interest to stay within the microscope field of view during frequent pressure perturbations, up to 20 times/s, during pressure ramp-up.

A similar solution was reported by Tanyeri and Schroeder,24 who used hydrodynamic trapping to precisely control the position of a bead via an image-based proportional-only controller. However, due to strand deformation, low SNR, limited pressure resolution, and stability, the requirements in this application differ substantially.

To minimise the residence time of a selected DNA molecule in the microchannel, the molecule should be kept as close to the entrance as possible. With the 1 mbar resolution of the pressure controller, the pressure setpoint will always be offset by up to 0.5 mbar from value needed for stationarity at the entrance. This causes oscillations around the target position, increasing average absolute displacement, and thus duration of the insertion phase.

Clearly, the average displacement distance could be reduced with better pressure resolution, a limitation that we resolved by developing and integrating algorithms for pulse width modulation (PWM) of the pressure set-point. This allowed a selected molecule to remain stationarily at the entrance during transport from microchannel to nanoslit, as demonstrated on a monodispersed sample of genomic DNA (i.e., T4 phage DNA at 50 ng/ml labelled with YOYO-1). Whereas PID is used to position a molecule at the nanoslit entrance, PWM is used to reduce the positional fluctuations around the entrance by increasing the effective pressure resolution. Thus, in order to probe the effect of PWM rather than PID, PID-stabilisation was enabled for 60 s to ensure steady-state near the nanoslit entrance, followed by a 60 second period of recording with or without PWM. This was repeated for several molecules, and the spatial offset of each molecule was recorded. As shown in Figure 5(a), PID by itself is sufficient for rough positional control, with 99.9% of all observations being within ±5 µm from the target setpoint and 44.0% of observations being within ±1 µm (∼6 pixels) from the setpoint. However, PWM further improves the stabilisation performance, with 92.3% of observations being within ±1 µm of the setpoint (Figure 5(b)).

To detect an insertion event, an interrogation area in the microchannel and one in the nanoslit (both 128 × 192 px) were compared during the real-time experiment. When a molecule is transported from microchannel to nanoslit, the criterion \( \langle I \rangle_{\text{microchannel}} < \langle I \rangle_{\text{nanoslit}} \) is fulfilled, marking the end of the insertion phase. The system then automatically sets all pressure setpoints to zero, translates the stage to the nanoslit, and adjusts focus via an autofocus algorithm.

A total of 25 focus estimation methods were benchmarked on micrographs acquired on our setup, using the MATLAB-toolbox distributed by Pertuz et al.25 The Tenengrad variance method was chosen for its accuracy, speed, ease of integration, and versatility, as it was found to perform well in both brightfield and fluorescence conditions. Upon completion of autofocus, the system automatically advanced to the next and final phase of the experiment, namely, the stretching phase.

As argued here, stabilisation via PID and PWM optimises the duration of the insertion phase, increases system throughput, and reduces risk of photo-nicking. The concepts presented here should be generally applicable to a range of microfluidics and/or fluorescence applications.

D. Phase 3: Stretching

In the final phase of the experiment, a DNA strand has been introduced into a nanoslit and now needs to be stretched along the nanoslit ending up with each end in opposite microchannels. During this process, we face a different set of challenges than in microchannels, which are general to applications of fluorescence microscopy in nanofluidics systems.

The random starting conformation of the DNA strand makes the shearing asymmetric. A flow is induced in the nanoslit which begins to shear the DNA molecule, but as the flow in each arm of the nanofluidic cross depends on the channel depth cubed, flows may also not be entirely symmetric due to fabrication tolerances. As a result, the Stokes drag pulling on the molecule in the cross is asymmetric, which may cause the strand to be pushed out of the nanoslit rather than elongated. To avoid this, the endpoints are constantly tracked by translating back and forth between the last known endpoint coordinates, and when asymmetric elongation is detected, a corrective pressure is applied. Full molecular extension is achieved when the endpoint positions stagnate, and the DNA may be imaged to extract a fluorescence image of the entire molecule. Then, the strand may be flushed out of the nanoslit to a waste outlet or to a collection outlet for further analysis via complementary methods, and the protocol can be repeated from phase 1.

![Without PWM](image1.png) ![With PWM](image2.png)

FIG. 5. Stabilisation performance in two control schemes. PID action alone (a) is sufficient to maintain a rough position within a few micrometers from a given setpoint. However, the addition of PWM (b) ensures that almost all observations are within 1 µm from the setpoint. Based on N = 2752 micrographs.
The shallow depth of the nanoslit prohibits the DNA molecule from bending out-of-plane, but the nanoslit width allows bending in-plane. In defect-free chips, which do not contain pinning sites, ensuring symmetric elongation is generally sufficient to resolve such coiling within the nanoslit.

The shallow depth of the nanoslit also increases stretching of the DNA, thus limiting the fluorescence signal. On the other hand, the higher hydraulic resistance of nanofluidics causes lower movement velocities. This enables longer exposure times or frame-averaging, which further increases the SNR as shown in Figure 3b.

Because the molecule is at least five times longer than the field of view when fully elongated, the stage must be translated in order to track molecule endpoints. This introduces the challenge of low-SNR machine vision on varying backgrounds. Despite the chip having been thoroughly bleached along the nanoslit prior to the experiment, a perfectly homogeneous background over the 440 µm nanoslit was not achieved, complicating background subtraction. Our solutions were to record a background map of the entire nanoslit before introducing DNA, by piecing together images taken along the nanoslit. Then, for each position visited during endpoint-tracking, the relevant part of this background map was excised and subtracted from incoming video frames, effectively correcting for background variations.

Once the molecule spans the nanoslit, the 440 µm of the molecule residing within it is fully elongated, and this part of the DNA may then be imaged by capturing frames along the nanoslit. The precise method for imparting a sequence-dependent melting pattern on the strand, in order for the optical readout to carry information, and the extraction of this information have already been described elsewhere.19

IV. CONCLUSION

In this study, we have automated the critical steps of the operation of a nanofluidic device for stretching and imaging megabase-long DNA molecules, thus demonstrating that manual manipulation of single DNA molecules can be avoided. We have shown that machine vision can be implemented to manipulate single DNA molecules in pressure-driven flows. In particular, we report our solution for using imaging of genomic DNA molecules as feedback for automation, despite the fact that both the shape and intensity of individual molecules change due to the hydrodynamic shearing in a micro and nanofluidic environment. The automation through machine vision of single molecules is a step toward improving the throughput of this unique platform for imaging fluorescence barcodes on megabase-long DNA.

ACKNOWLEDGMENTS

The authors gratefully acknowledge funding from the European Commission under the Seventh Framework Programme (No. FP7/2007-2013) under Grant Agreement No. 278204 (Cellomatic) and from the Danish Council for Strategic Research, DSF, under Grant Agreement No. 10-092322 (PolyNano). Furthermore, Marco Matteucci and Peter Friis Østergaard are acknowledged for kindly fabricating the polymer devices.

21See supplementary material at http://dx.doi.org/10.1063/1.492068 for a time-lapse of a DNA molecule selection and insertion.