Lab on a Chip

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/loc

Single molecule DNA intercalation in continuous homogenous elongational flow

Joshua W. Griffis, Mikhail M. Safranovitch, Shilpi P. Vyas, Sean Gerrin, Ekaterina Protozanova, Gene Malkin, ,Robert H. Meltzer*

PathoGenetix, Woburn, MA 01801, USA

*To whom correspondence should be addressed. Tel: 1-781-939-6481; Fax: (781) 938-0060; Email: robmeltzer1@gmail.com

The authors wish it to be known that, in their opinion, the first three authors should be regarded as joint First Authors.

ABSTRACT

Sequence-nonspecific staining of DNA with intercalating fluorophores is required for fluorescence-based length estimation of elongated DNA in optical mapping techniques. However, the observed length of a DNA molecule is affected by the relative concentrations of DNA and dye. In some applications, predetermination of DNA concentration may not be possible. Here we present a microfluidic approach in which individual DNA molecules are entrained by converging laminar sheath flows containing the intercalating dye PO-PRO-1. This provides uniform staining regardless of DNA concentration, and uniform elastic stretching of DNA in continuous elongational flow. On-chip intercalation provides a unique process for concentration-independent staining of long DNA fragments for the optical mapping method Genome Sequence Scanning (GSS), and normalizes intramolecular elasticity across a broad range of molecule lengths. These advances permit accurate mapping of observed molecules to sequence derived templates, thus improving detection of complex bacterial mixtures using GSS.

INTRODUCTION

Numerous biophysical investigations have studied the mechanical properties of DNA, and have described the structure of the polymer,[1-3] its interaction with ligands,[4-6] and its elasticity under tension. [1, 2, 5, 7] Several methods have been applied to stretch individual molecules of DNA, including

steric confinement in nanometer-scale channels,[8] adhesion to charged substrates,[9, 10] linear extension of individual tethered molecules using optical or magnetic tweezers [1, 2, 5, 7, 11] or entrainment of DNA in fluidic [12-16] or electrokinetic [17, 18] elongational flows We have previously reported a pseudo-two-dimensional microfluidic funnel that permits high throughput DNA stretching with a constant applied tension profile during analysis of individual DNA molecules in continuous homogenous elongational flow.[19]

The contour length of non-intercalated DNA may be determined by the separation of conjugated beads when force is applied using optical or magnetic tweezers,[20, 21] or by atomic force microscopy of molecules adhered to a substrate.[22] Alternatively, or in conjunction with the above methods, fluorescent intercalating dyes may be used to permit optical characterization of the molecule length. These fluorophores bind to DNA by inserting a planar aromatic moiety between adjacent basepairs.[23] Bis-cyanine dyes such as YOYO-1 or POPO-1 are commonly used for single molecule analysis because of their high affinity and 1000-fold enhancement of fluorescence quantum yield upon binding to DNA.[24, 25] Use of such probes, however, complicates analysis since intercalation of DNA changes the elasticity of the polymer, effectively lengthening [26-28] and untwisting the helix. [28-30] Both mechanical properties of DNA as well as interaction with intercalating dyes are affected by temperature, buffer ionic strength, pH, and the relative concentrations of DNA binding sites and dye. Achieving reproducible DNA elongation under tension, therefore, requires careful control of intercalation conditions. Additionally, suboptimal intercalation conditions can cause adhesion to glassware, crosslinking of DNA molecules and precipitation of DNA from solution.

In this study, we achieved uniform DNA staining conditions using the mono-intercalating precursor of POPO-1, PO-PRO-1 by introducing the fluorophore in converging sheathing flows upstream of a DNA stretching fluidic microfunnel. In this configuration, one molecule of DNA is introduced to the intercalating dye at a time, thus eliminating the concentration dependence on DNA and reducing the reaction kinetics to first order. The degree of DNA intercalation, therefore, depends only on PO-PRO-1 concentration.

The tension distribution along an extended molecule of DNA in the detection microfunnel is not uniform; rather it achieves a parabolic distribution.[13, 19] Maximum tension occurs at the center of the linearized polymer and decreases to zero at the ends. Each molecule, therefore, experiences a spectrum

of intramolecular tensions. This tension distribution is evidenced by the distribution of fluorescent sequence specific probes bound to the elongated DNA, effectively an intramolecular ruler with known sequence dependant markers. By using experimentally determined parameters for the elasticity of intercalated DNA, fluorescence traces of these sequence-specific probes can be aligned with high fidelity to templates predicted from DNA sequence. This capability forms the basis of an important application for high-throughput DNA stretching and analysis – Genome Sequence Scanning (GSS).

GSS is a platform technology well suited for bacterial identification and strain typing.[31-33] Detailed understanding of the elastic properties of DNA in elongational flow permits strain-specific identification of bacterial genomic DNA using either sequence-derived or experimentally generated reference templates. As each observed molecule is compared to reference templates individually, complex bacterial mixtures, with minor populations present at ~1% of the total population can be analyzed. These capabilities make GSS an important emerging tool for bacterial identification in food safety testing, outbreak epidemiology, and clinical diagnostics.[34-36]

MATERIALS AND METHODS

Reagents

Tris EDTA buffer (TE, 10 mM Tris, 1 mM EDTA, pH 8.0) was purchased as 100x stock from Sigma Aldrich (St. Louis, MO), and was diluted and supplemented with 0.025% NaN_3 prior to use. The intercalating dyes used in this study were POPO-1 and PO-PRO-1 (Fig. 2A,B, Life Technologies, Grand Island, NY). These dyes were selected for their high quantum yield, fluorescence enhancement upon binding DNA, and spectral separation from ATTO550 and ATTO647N, which are used for sequence-specific labeling (see below). POPO-1 has been used extensively in prior GSS investigations.[19, 31-33, 35-38] POPO-1 and PO-PRO-1 were diluted in DMSO to 10 or 100 µM stock concentration, and further diluted fresh to working concentration in TE.

Bacterial culture and DNA preparation

The following strains of bacteria were used in this study: *Escherichia coli* K12 (MG1655, GenBank U00096.2) [39, 40], *Escherichia coli* Sakai (ATCC BAA-460), *Salmonella enterica* LT2 (ATCC 700720)[41], and *Citrobacter koserii* (ATCC BAA895). 10^8 bacterial cells of overnight culture were loaded onto a custom

mini-reactor. [31] DNA was extracted using chemical and enzymatic reagents as described previously. [31, 32] Isolated genomic DNA was digested using either NotI or SpeI restriction enzymes (New England BioLabs, Ipswich, MA), generating a series of fragment lengths from 100 kb to 400 kb spanning the analyzable range of GSS (See ESI Fig S1). [19, 35] The digested DNA was then labeled using PNA fluorescent probes (PNA Bio, Thousand Oaks, CA) which bind to to 5'-GAAGAAAA-3' and 5'-GAAAAAGA-3' sequences on DNA (see 30 for detailed structure of the probes). The prepared DNA was eluted from the reactor at 1.0 – 1.5 ng/µl (as quantified by ethidium bromide-stained agarose gel) [32] and used for experiments. For some experiments, DNA was further diluted using TE.

Microfluidics

Two microfluidic DNA stretching devices were used in this study, previously described as CS 30 and CS 50.[19] Each device incorporates a high-throughput, continuous flow, constant strain-rate detection funnel, which constrains and accelerates the flow of DNA injected between two adjacent sheathing buffers (Fig. 1B). The two devices differ in their funnel taper geometry, which dictates the fluidic strain rate ($\dot{\varepsilon}$), and tension along a molecule of a given length at a specified fluid velocity. Details of the geometry differences and strain rates are provided in Table 1. Microfluidic devices were designed by PathoGenetix (Woburn, MA), and manufactured by Innovative Micro Technologies (IMT, Santa Barbara, CA). Microfluidic devices were fabricated in fused silica and etched to 2 µm depth, as described previously .[19]

Intercalation

We investigated two methods for reacting DNA with intercalating dyes. In-tube intercalation, as used in previous studies, permitted sufficient time for intercalating dye to equilibrate with DNA. DNA was pre-mixed with dye in a microcentrifuge tube, allowed to incubate for several minutes, and loaded to the DNA injection port of the microfluidic device. Alternatively, using on-chip intercalation, unstained DNA was loaded to the injection port and dilute intercalating dye (50 – 500 nM) was introduced to the sheathing buffer streams of the chip. DNA was therefore kept separate from the intercalator until the fluid streams converged at the union with the stretching funnel. From fluidic models, the total reaction time was limited to approximately 300 ms.

The two methods of intercalation required different metrics for quantifying the amount of dye added. For in-tube intercalation, the molar ratio of dye to DNA base pairs was reported as the intercalation ratio [POPO-1]:[DNA bp]. Prior to intercalation, DNA concentration was quantified by ethidium-stained agarose gel, and an appropriate dilution of intercalator was added to achieve the desired concentration ratio. For on-chip intercalation, fixed concentrations of dye were added directly to the device, with no required DNA pre-quantitation. Only the intercalator concentration was therefore reported.

Microscopy

In order to visualize DNA intercalation and interaction of stained DNA with the microfluidic device, the chip was imaged using an inverted fluorescence microscope Eclipse TE 2000-S with a 10x Pan Fluor objective, 0.74 NA (Nikon, Tokyo, Japan). POPO-1 and PO-PRO-1 fluorescence was imaged through a CFP-2343A-NTE-zero filter (Semrock, Inc., Rochester, NY) and a SensiCam qe CCD camera (Cooke Corporation, Romulus, MI).[35] Images were processed and false-color filters were applied for contrast using ImageJ software.[42]

Detection

DNA detection in fluidic microfunnels has been described previously.[13, 19] Individual fluorophore-labeled DNA molecules elongate in the continuous accelerating flow along the funnel, and pass over a series of excitation spots of laser light (Fig. 1C). Two 445 nm excitation spots, separated by 20 µm along the detection funnel excite the intercalator bound to DNA and provide velocity and length information (Fig. 1B). 532 nm and 640 nm laser light spots, positioned 5 and 10 µm upstream of the first 445 nm spot excite fluorophores attached to PNA probes. Fluorescence intensities are reported in arbitrary units. Within a single experiment, the fluorescence intensities between molecules are reproducible. Fluorescence may vary between experiments due to optical alignment, laser power, and alignment of the microfluidic chip to the detector. Quantitation between experiments is therefore less reliable than analysis of relative fluorescence intensity within a single experiment.

Data Analysis – Molecule length and average intercalator signal intensity

The velocity and length of each molecule was determined from the time of arrival and duration of transition through two 445 nm detection spots separated by a 20 μm distance using the fluorescent signals emitted by from the intercalating dye. As shown in Fig. 1D, the average intercalator signal was plotted versus the observed length for all molecules in a single experiment. Each cluster in this heat map corresponds to a collection of DNA fragments with similar observed lengths. As has been shown previously, plots of this sort correlate well with molecule length discrimination observed by pulsed field gel electrophoresis.[32]

Data analysis – Clustering of similar molecules

As in gel electrophoresis, it is not possible to differentiate DNA fragments with very similar lengths purely by the observed intercalator fluorescence signal. The presence of sequence specific probes on the DNA distinguishes fragments of similar lengths based on their unique fluorescent signatures determined by underling sequence.[35, 36] A previously described clustering algorithm was used to determine averaged elastic properties of individual restriction fragments of known polymer length *L,*(reported in kb). [36] This algorithm compares each molecule to all others within a selected length range and forms clusters of molecules with statistically similar fluorescence traces from their site-specific probes. In this analysis, each restriction fragment is expected to generate two mirror image clusters, representing forward and reversed orientation molecules, as each molecule has equal probability of entering the stretching funnel in forward or reversed orientation. The extension (*X,* in μm) for each of these clusters can then be assigned to their known restriction digest fragment lengths (in kb). The contour length of a DNA fragment was calculated as *X/L* (μm/kb, or nm/bp).

The experimentally determined average templates can further be used to populate a database for single molecule template based classification. Averaged templates corresponding to *E.coli* K12 NotI fragments 248.5, 250.1, and 250.5 kb in length are shown in Fig. 1E. These fragments are not resolvable by length analysis alone, but are clearly identified by site-specific tagging patterns.

Data Analysis – Template-based single molecule classification

Template based molecule classification is based on a calculation of likelihood that each molecule could originate from one of a multitude of representative templates in a database. [35] These templates may be generated either from clustering experimental data as described above, or from computational predictions derived from whole genome sequence. The likelihood calculation is based on the statistical model of expected distribution of photons along a molecule. The assumed Poisson-Gamma distribution accounts for experimental noise and stochastic events at the single fluorophore level.

Each molecule in a data set is compared to every template in the database, and is assigned to a single best fit template. Statistics indicating the quality of fit (average log-likelihood, *αLL*) and relative number of assigned molecules from the total number of molecules processed in a single experimental sample (% fraction) are derived for individual fragment templates, or summed for all representative fragments for a single organism. Ambiguous molecules, either due to poor site-specific tagging or excessive stochastic noise, may be assigned randomly to any template, thus establishing a probability for false positive identification. Randomized null templates are therefore added to the database in order to establish the false positive frequency. True positive identification requires population of a single template with higher log-likelihood and greater % fraction above this false positive threshold.

RESULTS

Convergent stream intercalation produces consistent, concentration-independent DNA stretching in elongational flow

Intercalation of purified DNA using the bis-intercalating dye POPO-1 is very sensitive to reaction conditions. In particular, the relative concentration of dye to DNA base pairs affects the availability and analyzability of DNA in continuous elongational flow microfluidics, as shown in ESI Fig. S2. When DNA was present in high excess of intercalator, [POPO-1]:[DNA bp] < 0.1, staining of DNA was inadequate for reliable fluorescence detection. When the POPO-1 concentration exceeded the concentration of DNA ([POPO-1]:[DNA bp] > 1:1), DNA precipitated from solution and could not be reliably transferred to the fluidic device for stretching and detection. For the range of concentration ratios between these extremes, DNA could be detected, but the observed molecule lengths varied significantly, depending on the intercalator concentration. This variation in DNA stretching efficiency hinders DNA length

estimation in further analysis. In previous publications, the ratio of intercalator to DNA was fixed experimentally to a value between 1:3 and 1:5 to minimize the impact of intercalation on DNA stretching reproducibility.[19, 31, 32, 36]

In order to normalize DNA stretching efficiency regardless of DNA concentration, intercalating dye was introduced to the sheath buffer ports of the microfluidic device (Fig. 1A). Initial experiments using POPO-1 as the intercalating dye were encouraging: intercalator interacted with DNA on the fly and fluorescent signal was observed despite the short reaction time (~300 ms) allowed in the stretching fluidic channel. However, after 10-15 minutes of operation detection of molecules in the channel ceased. Fluorescence microscopy of the junction between the DNA injection channel and the sheathing buffer channels revealed that the microfluidic device was reproducibly occluded by a mat of intercalated DNA (Fig. 2A). This occlusion could be dissociated by copious flushing of the device with 100 mM NaOH. Most likely DNA sticking was due to the short time of the interaction of the dye with DNA such that only one aromatic moiety of POPO-1 intercalated the double helix while the other end remained free in the solution able to bind to other DNA molecules and glass surfaces.

The intercalation reaction was repeated using the monomeric dye, PO-PRO-1 in an attempt to eliminate crosslinking of DNA. An initial test fluorescence binding assay suggested that intercalation approached saturation near 250 nM PO-PRO-1 in the presence of 50 nM DNA base pairs. At concentrations above 1 μM, the observed fluorescence intensity appeared to decrease, either due to intercalator-induced autoquenching of the fluorescent signal, or due to aggregation and precipitation of the stained DNA (ESI Fig. S3). [43] For initial on-chip intercalation, we therefore used 250 nM PO-PRO1 in low ionic strength TE buffer (0 NaCl) in order to accentuate electrostatic interactions between the dye and DNA. Fluorescence microscopy revealed the presence of a single stream of intercalated DNA emerging from the DNA injection channel (Fig. 2b). The stream of DNA achieved maximal brightness across the entire DNA stream width within 50 μm of the DNA injection channel, indicating rapid association of DNA with dye (note that the total channel length is 1450 μm for CS 50). No aggregation or surface adhesion was observed, even after several days of continuous flow. Use of the mono-intercalating dye was therefore necessary to promote successful on-chip intercalation by eliminating DNA crosslinking and adhesion to the chip surfaces.

The extensibility of a DNA molecule intercalated in tube with POPO-1 depends on the ([POPO-1]:[DNA bp] ratio. Serial dilutions of DNA were incubated with 1 µM POPO-1 prior to loading onto a stretching funnel for 30 minute data acquisition. The contour length L_C, or ratio of observed molecule length (X) to known length derived from sequence (L) was determined for the 250.5 kb fragment of *E.coli* K12 NotI digest as a common metric for all data sets. We observed that L_C increased as the ratio of intercalator to DNA concentration increased (Fig. 2C). Only a ten-fold range of intercalator to DNA ratio was achieved, as at low dilutions of DNA the frequency of molecules reaching the detector dropped precipitously (corresponding to the observed onset of fluorescence quenching and precipitation, see ESI Fig S2, S3).

In contrast, when similar DNA dilutions were intercalated on the chip with 250 nM PO-PRO-1 in the sheathing buffer streams, a uniform elongation of the 250.5 kb fragment was observed across 32-fold dilution series of DNA (0.04 to 1.5 ng/µl DNA) (Fig. 2D). Through on-chip intercalation with PO-PRO-1 therefore we achieved the primary goal of eliminating DNA concentration dependence for continuous flow stretching fluidics.

DNA extensibility varies with intercalator concentration

The extent of stretching of a DNA molecule in elongational flow microfunnels depends on the tension applied to the molecule and the elasticity of the polymer. Tension along the length of a molecule ($T(x)$) has been described previously as:

$$T(x) = \zeta \int_{x}^{x_{head}} (v(s) - v_{mol})ds, \qquad \text{Eq. 1}$$

where s is the distance along the length of the extended molecule, $v(s)$ is the fluid velocity along the central streamline at distance s, v_{mol} is the molecule velocity (average of the fluid velocity along the length of the extended molecule, x is the position within the stretching funnel, and ζ is the molecular drag coefficient. ζ has been previously estimated in computational models and confirmed experimentally to be 0.61 cP.[19, 44] This results in a parabolic tension profile along the length of the molecule, where tension is maximal at the center of the molecule, and approaches zero at the ends. The peak tension T_{max} along a molecule of extension X is approximately:

$$T_{\max} = \frac{1}{8}\zeta X^2 \dot{\varepsilon}(x), \qquad \text{Eq. 2}$$

where $\dot{\varepsilon}$ is the uniform strain rate in the detection funnel defined by the funnel geometry and fluid velocity.[19] The average of this parabolic tension distribution (T_{avg}) is 2/3 T_{max}.

To demonstrate the effect of varying $\dot{\varepsilon}$ on DNA elongation, *E.coli* K12 NotI digest was processed on two stretching funnels designed for different strain rates at 50 µm/ms fluid velocity (0.065, 0.1003 ms^{-1}), using convergent stream on-chip intercalation with 250 nM PO-PRO-1. Sixteen individual restriction fragments were identified by clustering and plotted in Fig. 3A versus their expected lengths. Under these conditions, stretching at the higher strain rate device resulted in greater elongation of the DNA than on the lower strain rate device. The effect was most apparent for $L >$ 200 kb.

It is convenient to compare the elongation of multiple populations of DNA fragments by normalizing to their contour length L_c. These values were determined for all fragments from Fig. 3A, and plotted versus maximal tension in Fig 3B. Maximum tension values were calculated by Eq. 2 using the observed fragment length X and $\dot{\varepsilon}$ specific to either the CS 30 or CS 50 device driven at 50 µm/ms as presented in Table 1. Fig. 3B shows that across a broad range of T_{max}, DNA elasticity behaved consistently and predictably. Below ~100 pN T_{max}, there was an approximately linear relationship between tension and elongation. Above this threshold, elongation became increasingly nonlinear. The behavior in the CS 30 device superimposed with that of the CS 50 device, confirming that the elongation scaled with peak tension. Previous studies have demonstrated that addition of intercalating dyes such as ethidium or YO-PRO-1 increase the elasticity of DNA and can eliminate the overstretching threshold exhibited by unstained DNA at ~65 pN, agreeing with our observations.[20] For further experiments, we limited analysis to DNA molecules held to ~100 pN T_{max} in order to remain within the approximately linear tension/elongation regime. It should be noted that 100 pN T_{max} is equivalent to 66 pN average tension along the molecule. L_c under these conditions ranged from 0.37 nm/bp to >0.5 nm/bp. These values were greater than the expected behavior of non-intercalated DNA where L_c = 0.34 nm/bp, in accordance with molecular structure. [45]

L_c values for *E.coli* K12 NotI digest fragments were plotted vs. T_{max} for experiments spanning a range of PO-PRO-1 concentrations from 50 nM to 500 nM (Fig. 4A). The extension ratio appeared

proportional to the peak tension applied to each molecule. The slope of these plots did not vary with intercalator concentration, but the projected zero-tension elongation increased with higher intercalator concentrations (Fig 4B). In contrast, DNA incubated with different relative concentrations of POPO-1 (in-tube intercalation) varied in both slope and intercept (Fig. 4C,D).

The tension versus extension characteristic of the DNA may be approximated using a simple bead and spring model. In such a model, the polymer may be represented as a series of discrete segments connected by extensible springs. The slope of the tension-extension graph (*k,* Fig. 4 B, D, filled circles) effectively yields the average spring constant for the connecting springs, and the ordinate intercept, effectively the zero-tension extension condition, yields the inter-segment spacing (*b*, Fig. 4 B,D open circles). Note that the latter value could not be measured directly, but was extrapolated from the fitted curves in Fig. 4 A, C. In the case of POPO-1 intercalation (Fig. 4 D), both the slope and intercept varied with increasing concentration of intercalator. As the bis-intercalator physically cross links adjoining DNA base pairs, it follows that the elasticity of the polymer depends on the relative proportion of cross linked vs. non-cross linked DNA segments.

In contrast, PO-PRO-1 intercalation cannot bridge adjoining DNA base pairs, and therefore the extensibility of the DNA reflects the inherent elasticity of the polymer, within the tested range of intercalator concentrations (Fig. 4B). The extension of the polymer under zero tension conditions, however, was affected by the population of intercalator-bound sites because addition of intercalator extends and untwists the DNA polymer.[28] Maximal zero-tension DNA extension would be expected to saturate as all binding sites become occupied with intercalator at high intercalator concentrations. However, this was not demonstrated below 500 nM PO-PRO-1. Above 500 nM PO-PRO-1, significant background fluorescence in the intercalator detection channels precluded experimental data acquisition. At the other extreme, the minimal intercalation-induced extension would be expected at zero PO-PRO-1 concentration. This was also not observable in this experimental configuration, but was extrapolated by a linear fit to the collected data (Fig. 4B). Within experimental error, we determined that the minimal contour length was 0.34 nm per base pair, in agreement with the classical estimate of base pair pitch.[45]

Intramolecular polymer elongation with a parabolic tension distribution

Numerical models of an individual molecule in the constant strain rate detection channel indicate that the tension along its length assumes a parabolic distribution.[19] We have previously demonstrated that the elasticity of DNA, as measured by the overall length of a molecule, depends on the applied tension, as well as the concentration of intercalator available (Fig. 4). As tension along the molecule is not uniform, the extension of each segment along the length of the molecule will vary with its applied tension.

Sequence specific fluorescent labeled probes that decorate multiple sites on long molecules of DNA provide a convenient tool to investigate intramolecular elasticity in the presence of non-uniform intramolecular tension. Because molecules are under constant acceleration throughout their passage through the detection channel, the length of DNA (in μm) encompassed within a single bin of data varies along the length of the molecule, and also varies across different molecule lengths. A first-principle solution for correcting for acceleration has previously been presented.[19] Similarly, because of the tension distribution along the molecule, the polymer length (in kb) encompassed in each micron of DNA varies along the length of each molecule. A two-stage correction is therefore required to convert data collected in the time-domain, first to the micron-domain, and secondly into the polymer length-domain.

The effect of non-uniform elasticity of the DNA molecule can be seen in Fig. 5A, which represents experimental average optical traces obtained for molecules of 126, 217, and 291 kb length from the SpeI digest of *Salmonella.* The magenta traces correspond to the distribution of site selective PNA labels designed to tag 8-bp long sites with 5'-GA AGAAAA-3' sequence on DNA. These traces are superimposed on sequence-derived linear templates, corrected for acceleration as described previously [19], but without elasticity correction (blue traces). A significant shift in the location of peaks relative to the linear templates is apparent, most notable near ¼ and ¾ along the length of the trace. An empirical correction for elasticity can be achieved by fitting a simple sinusoidal function to the abscissa of these plots,

$$x' = A\sin\left(\frac{2\pi x}{X}\right) \qquad \text{Eq. 4}$$

where *x'* is the rescaled length axis, *A* is the shift amplitude, and *x* is the position along the observed length of the molecule *X*. The required shift amplitude can be determined experimentally by visual

matching of the plots on Fig. 5A for molecules of different lengths and for samples processed with different intercalator conditions.

Elasticity correction was experimentally determined for DNA ranging in length from 40 to 140 µm from *Salmonella* SpeI digest processed by in-tube POPO-1 intercalation at a targeted 1:3 [POPO-1]:[DNA bp] ratio and by using convergent flow with 250 nM PO-PRO-1. The required shift amplitude (*A,* Eq. 4) was plotted vs. peak molecule tension (Fig. 5D). PO-PRO-1 intercalation data demonstrated a proportional relationship between intramolecular tension and required shift amplitude. In contrast, the required shift for DNA labeled with POPO-1 was not proportional to molecule tension. The POPO-1 intercalated DNA appeared to be less elastic for tensions below 50 pN, compared to PO-PRO-1 intercalated DNA. Above this threshold, however, DNA became more elastic. The elasticity correction for POPO-1 intercalated DNA was also less uniform than that observed for PO-PRO-1 stained DNA. This was likely due to the inherent sensitivity of the DNA staining reaction to the ratio of DNA to intercalator which might have varied in different experiments because of DNA quantification error.

The uniform relationship between experimental elasticity correction and peak molecule tension observed for PO-PRO-1 intercalated DNA presented an opportunity for deriving a model for intramolecular DNA elongation under parabolic tension distribution. As described above, in a constant strain rate detection channel, the DNA molecule moves at the average of the fluid velocity along its length, and a parabolic tension distribution is established along the length of the molecule. From experiments describing the full molecular elongation with increasing intercalator concentration, we observed that the elongation *vs.* tension relationship had a uniform slope, independent of intercalator concentration (Fig. 4B). The ordinate intercept of this relationship, however scaled linearly with intercalator concentration. From these observations, we derived parameters for the intrinsic DNA extensibility (k = 0.0011 µm/kb pN) and extensibility of DNA with 250 nM intercalator (b = 0.37 µm/kb). Solving for the number of kb encompassed in a single µm, for DNA held under tension T_{max} becomes:

$$\frac{L}{X} = \frac{1}{kT_{\max} + b} \qquad \text{Eq. 5}$$

The polymer length distribution along a molecule of a given observed length becomes non-linear with this distribution – more kb are encompassed at the ends of the molecule compared to the

center (Fig. 5B, solid line). This is distinct from a linear tension distribution model, with uniform distribution of polymer length (dashed line). Subtracting the parabolic tension model from the linear model results in a pseudo-sinusoidal "correction curve" (Fig 5C, solid curve). In fact, this relationship can be best fit with a third order polynomial equation, rather than the sinusoidal approximation (Eq. 4; Fig 5C, dashed line). The amplitude of this fitted curve, using only derived parameters is plotted in comparison to the experimental best fit data for the *Salmonella* example (Fig. 5D, open circles). The derived elasticity correction is in excellent agreement with experimentally determined elasticity parameters (Fig. 5D, red filled circles). Such modeling is not possible for POPO-1 intercalation, due to the non-uniform elasticity of the DNA across a range of molecule tensions (Fig. 5D, blue squares).

Correction for intramolecular elasticity permits detection of multiple bacterial species in complex mixtures.

In GSS, fluorescent traces from individual molecules are compared to a database of representative fluorescence templates.[35] Such templates may originate from experimental data, using the clustering algorithm described above. Alternatively, templates may be generated computationally from whole genome sequence; such templates are "linear" in that polymer length is proportional to genomic content. Elasticity correction must be applied to experimental templates in order to standardize the template database for molecule classification.

Fig. 6A shows the results of classifying molecules from an *E.coli* Sakai SpeI digest against a database consisting only of sequence-derived templates. 34903 molecules were classified. The confidence of detection for each fragment (*aLL,* see Methods) is plotted against the fractional population of all molecules assigned to each fragment. Sequence-derived templates comprising the database were tested with and without elasticity correction. When elasticity correction was applied, each expected fragment was detected with greater *aLL* and higher number of assigned molecules. The average *aLL* across all fragments using elasticity correction was 5.9 ± 2.7, and 21171 molecules (52.2%, sum of all detected fragments) were confidently assigned to the 17 representative templates. In contrast, without elasticity correction, average *aLL* across all fragments dropped to 2.6 ± 0.6, with only 5996 correctly assigned molecules (17.3%).

Molecules not correctly assigned to their target templates typically failed due to ambiguous fluorescence traces caused by inadequate tagging, overlapping fragments, or poor stretching. These molecules were assigned to best-fit templates in the database, usually with poor *aLL*. Randomized null templates were added to the database in order to establish a threshold for confident detection for such randomly assigned fragments (false-positive identification). Confident detection is established for a fragment if the product of *aLL* and fractional population (total log-likelihood, *tLL*) was greater than five standard deviations above the average *tLL* of null templates. This hyperbolic distribution is plotted in Fig 6A. One fragment from the data classified with non-corrected templates falls below this threshold (*tLL* = 0.011 ± 0.031).

The confidence of detection of an organism can be expressed as the sum of *tLL* for all expected fragments. Six replicate data sets of samples prepared from *E. coil* Sakai BAA460 were compared against multiple template sets (Fig. 6B). As described above, use of elasticity corrected, sequence-derived templates resulted in greater confidence of detection than when templates were not corrected. Templates may also be generated by processing of an isolated organism through GSS and clustering averaged templates from experimental data. No intrinsic elasticity correction is required in order to directly classify molecules against such unmanipulated templates, as both template generation and data acquisition is performed under identical stretching and intercalation conditions. However, such templates cannot be used in conjunction with sequence-derived templates. To reconcile these two template sources, clustered templates may be manipulated to remove elasticity, linearizing the templates or restoring a proportional kb/µm relationship. Direct comparison of experimental data to linearized experimental templates behaved similarly to non-corrected sequence-derived templates. Linearized experimental and sequence-derived templates can then be corrected for elasticity using identical manipulation to achieve highly confident detection. Use of this method also allows for conversion of templates generated under different intercalation or stretching conditions into a condition-independent format which can then be adjusted as needed for different experimental conditions. This allows for various stretching fluidics or intercalators to be evaluated without need to experimentally rebuild the testing database of reference templates.

Characterization of complex bacterial mixtures

As each molecule in a GSS dataset is compared to reference templates individually, complex mixtures of genomes may be classified simultaneously. One such experiment from three replicates is presented in Fig. 6C. A bacterial sample was prepared targeting 0.5-2% *E.coli* Sakai in the presence of excess *Citrobacter koseri*. The ratio of *E.coli* Sakai to *Citrobacter* was confirmed by plating to be 0.18 % – 0.7 % for different samples. *aLL* for *Citrobacter* and *E.coli* Sakai digested with SpeI is plotted against percent of classified molecules using elasticity corrected or non-corrected templates. *aLL* for both major and minor components was significantly higher for elasticity-corrected templates. This was particularly important for confident detection of the minor component *E.coli* Sakai. As shown in the inset, the majority of *E.coli* Sakai fragments failed to exceed the detection confidence threshold when elasticity correction was not applied. With appropriate compensation for elasticity, however, confident detection of the majority of fragments was achieved. Fig. 6D shows the number of detected above threshold fragments as a fraction of fragments expected for the digest for both the major and the minor component of the mixture. Across three replicate experiments, all expected fragments of *Citrobacter* were detected using elasticity correction, and this suffered slightly in one experiment without correction (paired T-test, p=0.42). There was a significant improvement, however, in detection coverage for *E. coli* (minor component) when elasticity correction was applied (paired T-test, p = 0.01). Appropriate compensation for elastic deformation of DNA stretched in GSS fluidics in the presence of intercalating dye is therefore critical to detection and classification of bacterial genomes. The quality of matching experimental data to templates for isolated bacteria is greatly improved with elasticity correction, and detection of minor components in mixed bacterial samples is not feasible without it.

DISCUSSION

DNA intercalating dyes are common molecular probes for a wide variety of biochemical and biophysical applications.[46] The spectrally diverse family of cyanine fluorophores are particularly useful because of their high binding affinity, low rate of dissociation, high quantum yield of the DNA-bound form, and substantial increase in fluorescence upon binding to DNA. [24, 25] Representative fluorophores include the dimeric dyes POPO-1, YOYO-1, TOTO-1 as well as their monomeric counterparts PO-PRO-1, YO-PRO-1, and TO-PRO-1. Their characteristics have made intercalating dyes from this family useful in gel electrophoresis,[25, 47] single molecule biophysical studies,[16] nuclear staining for physiological assays,[48, 49] flow cytometry,[24] real-time PCR,[50] and optical mapping of long genomic DNA.[51]

Intercalating dyes have a multitude of effects on the mechanical properties of elongated DNA. Binding YOYO-1 or YO-PRO-1 causes elongation [26-28] and untwisting of the double helix.[28-30] Notably, some intercalators appear to enhance DNA extensibility under tension and protect against overextension of DNA elongation at tensions above 65 pN, as has been observed for non-intercalated DNA.[5, 26, 52] These effects on DNA stretching depend on the ratio of intercalating dye to DNA binding site concentrations. Careful control of intercalation conditions is therefore required for accurate polymer length assessment.

We have developed a method for intercalation of one DNA molecule at a time in sheathing flows of the intercalator PO-PRO-1 in a microfluidic DNA stretching device. The extent of dye-induced elongation is therefore determined only by dye concentration, which can easily be controlled and standardized across experiments. On-chip intercalation therefore meets a fundamental requirement for automated, high-throughput screening using GSS, where pre-quantitation of DNA would be prohibitive.

Use of the mono-intercalating dye PO-PRO-1 also eliminated a significant difficulty with bis-intercalating dyes such as POPO-1, namely, crosslinking and precipitation of DNA. The binding mechanism of POPO-1 to DNA is complicated. Regardless of intercalator concentration, a significant number of molecules are bound to DNA in a mono-intercalated state; one aromatic moiety may be intercalated into the DNA, while the other may be free or bound to the DNA minor groove.[27, 53] These available charged aromatic groups may form additional bonds to the glass surface of the fused silica microchannel, or to additional DNA molecules flowing by, thus leading to precipitation and aggregation of the intercalated DNA. We hypothesize that these free charged moieties are not available when the monomeric dye PO-PRO-1 is bound to DNA, thus eliminating crosslinking and sticking. This allows for much longer effective life spans for the microfluidic device, and helps to eliminate carryover in sequential sample processing.

Use of the monomeric dye as opposed to the dimeric dye also affects the extensibility of molecules spanning a broad range of observed lengths. DNA stained with PO-PRO-1 appeared to have consistent elasticity regardless of the concentration of dye. Increased dye concentration, however, did contribute to a non-elastic pre-extension of the molecule. In contrast, the concentration of the dimeric dye POPO-1 affected both elastic and non-elastic elongation of the DNA. We rationalized this difference by assuming that increased POPO-1 concentration increases the opportunity for bis-intercalation, thus

linking adjoining DNA base pairs with additional molecular bonds. For the monomeric PO-PRO-1, all binding sites along the DNA may be occupied, but there is no mechanism for cross linking adjoining sites.

DNA stretching in continuous homogenous flow in GSS micro funnels provides a unique platform for investigating the interaction of DNA with intercalating dyes. Because of the mechanism in which tension is applied to the molecule, each molecule simultaneously experiences a spectrum of tensions. This is similar to the behavior in cross-slot fluidic devices,[12, 16] but distinct from that observed in optical tweezers or AFM experiments.[11] Unlike any of these methods, GSS has very high molecular throughput and efficiency. Aggregate statistical behavior of thousands of individual molecules can be compiled in short experimental acquisitions. GSS therefore provides a convenient platform for the biophysical characterization of the effects of tension on DNA binding ligands. A variety of experimental conditions including ionic strength, pH, temperature, reaction time, tension, and choice of fluorescent dye may be readily manipulated. Future studies will investigate the manipulation of such variables on DNA binding ligands in extensional flows.

Implementation of on-chip intercalation with PO-PRO-1 has important implications for the primary intended application of Genome Sequence Scanning – genomic comparison and strain specific identification of bacteria. Robust and reliable matching of site-specific tagging patterns to template traces is the fundamental goal of GSS. On-chip intercalation provides a uniform tension-elongation regime along the length of the molecule, as distinct from the non-uniform behavior observed previously with POPO-1 intercalation. Use of PO-PRO-1 permits predictive modeling of the affine transformation of tag locations (predicted from sequence length measured in kb) to observed locations (measured in microns). Better alignment of experimental and template data improves detection sensitivity, particularly for detection of rare targets in complex bacterial mixtures.[35] This capability subsequently improves the applicability of GSS in a wide range of studies, including biodefense monitoring,[35] clinical diagnostics,[36] and food safety testing.[34]

ACKNOWLEDGEMENT

The authors wish to thank Jimmy Symonds, and Maura Faggart for sample preparation. Deepak Ropireddy assisted greatly in literature review. Douglas Cameron, Donald Crothers, John Harris, and Rudolf Gilmanshin provided insightful conversations and critical review of this manuscript.

Tables:

Parameters	CS 30	CS 50
Stretching funnel length (μm)	600	1100
Detection channel length (μm)	350	350
F_1 (μm^3)*	4.5×10^6	1.8×10^7
F_2(μm^2) *	3433	7361
Etch depth (μm)	2	2
$\dot{\varepsilon}_{30}$ # (ms^{-1})	0.062	0.039
$\dot{\varepsilon}_{50}$ # (ms^{-1})	0.1033	0.065
$T_{max\ 30}$ (pN) †	47.2	29.7
$T_{max\ 50}$ (pN) †	78.8	47.3

Table 1: Functional parameters for GSS constant strain-rate funnels. * F1, F2: taper coefficient for stretching and detection funnels, respectively. # fluid strain rate achieved at 30 or 50 μm/ms center-stream fluid velocity. † Maximum tension achieved for 100 μm molecule at 30 or 50 μm/ms fluid velocity.

Figure legends:

Figure 1: Continuous flow single-molecule intercalation and detection. A) Diagram of DNA-stretching microfluidic chip. The fluidic channel consists of 1) DNA loading port, 2) sheathing buffer channels, 3) DNA stretching and interrogation channel, 4) vacuum/waste port. B) Cartoon of DNA extension in stretching fluidic channel and passing through four excitation laser spots. C) Fluorescence intensity traces for a single stretched DNA molecule. Blue, light blue: PO-PRO-1 signal; red, green: sequence-specific tag signal. D) DNA stretching heat map. *E.coli* K12 NotI digest was intercalated at 1:3:[POPO-1]:[DNA bp] ratio and stretched in CS50 chip. Average POPO-1 fluorescence intensity for each molecule is plotted against observed molecule length (*X*). Color scale of each pixel across 16 shades from grey to light blue corresponds to the number of molecules with the same intensity and length. This example comprises 200,260 individual molecules collected in a 30 minute acquisition, and each distinct cluster of

long molecules represents ~ 300 to ~ 2000 molecules, depending on the number of fragments of similar extended length present. E) Representative average fingerprints of 248.6, 250.2, and 250.5 kb fragments from the marked cluster in D). Traces for fluorescent tag binding to 5'-GAAGAAAA-3' sequence on DNA are shown.

Figure 2: DNA intercalation using POPO-1 (bis-intercalator) and PO-PRO-1 (mono-intercalator) A) On-chip intercalation using 250 nM POPO-1 in sheathing buffer caused adhesion of intercalated DNA to glass surfaces. B) On-chip intercalation using 250 nM PO-PRO-1 eliminated DNA aggregation and adhesion. C) Extension of 250.5 kb fragment form *E.coli* K12 NotI digest plotted for multiple [POPO-1]:[DNA bp] ratios (in-tube intercalation). Ratio of observed length to known polymer length plotted vs. [POPO-1]:[DNA kb]. D) Extension of 250.5 kb fragment from *E.coli* K12 NotI digest spanning 32-fold serial concentration dilutions using sheath buffer intercalation with 250 nM PO-PRO-1.

Figure 3: DNA elongation increases with applied tension. A) Observed length (*X*) plotted vs. known polymer length (*L*) for 16 fragments of *E.coli* K12 NotI digest run on high strain rate (CS 30, ●) and low strain rate (CS 50, ○) stretching funnels at 50 μm/ms fluid velocity. Single representative paired experiment. Each point represents average *X* for 200-500 molecules. B) L_c from experiment in A) plotted vs. maximum tension on the molecule. Error bars represent standard deviation of three replicate experiments.

Figure 4: Effect of intercalator concentration on DNA stretching. A) 11 fragments of *E.coli* K12 NotI digest stretched on low strain rate GSS chip at 50 μm/ms with increasing [PO-PRO-1] in the sheathing buffer (▲50, △100, ■150, □250, ●400, ○500 nM). B) Slope (*k,* right axis , ●) and intercept (*b*, ○, left axis) for linear extension vs. tension plots as shown in A) as a function of [PO-PRO-1]. C) Linear extension vs. tension plots for POPO-1 intercalation with increasing [POPO-1]:[DNA] (▲0.1, △ 0.16, ■

0.25, □ 0.4, ● 0.67, ○ 1.33). D) Slope (*k,* right axis, ●) and intercept (*b*, ○, left axis) for plots shown in C as function of [POPO-1]:[DNA].

Figure 5: Elasticity correction for parabolic tension distribution. A) The tension distribution along a DNA molecule is revealed in the spatial distribution of sequence specific labels. Representative traces for three fragments from *Salmonella* SpeI digest. Experimental data (magenta) is compared to sequence-derived templates (blue). Left-hand figures demonstrate mismatch in uncorrected experimental data. Right-hand figures demonstrate correct matching after sinusoidal correction. B) Theoretical model for DNA extention (kb/μm) distribution along a 100 μm molecule. Model assumes a parabolic tension distribution, and bulk elasticity for DNA intercalated with 250 nM PO-Pro-1 (solid line). In contrast, a model for DNA with uniform tension distribution is presented (dashed line). C) Shift amplitude as function of position along molecule length for tension derived model (solid line), as compared to sinusoidal approximation (dashed line). Sinusoidal approximation was used for simplified correction. D) Sinusoidal correction amplitude for *Salmonella* fragments measured using 250 nM PO-PRO-1 on-chip intercalation (red), plotted vs. maximal tension on the molecule (6 replicates, red circles). Fitted curve is linear. Similar data collected for 6 replicates, intercalated in-tube with 1:3 [POPO-1]:[DNA] (blue squares). Fitted curve is quadratic. Theoretical model for required shift amplitude, assuming 250 nM PO-PRO-1 is shown (open circles). Fitted parameters were: $\zeta = 0.61$ cP, $\varepsilon = 0.065$ ms^{-1}, $k = 0.0011$ μm/kb pN, $b = 0.37$ μm/kb.

Figure 6: Impact of DNA elasticity correction on single-molecule classification. A) *E.coli* Sakai SpeI digest was processed on low strain rate GSS fluidics at 50 μm/ms with 250 nM PO-PRO-1. Molecules classified against database consisting of sequence-derived templates with (●) and without (○) elasticity correction. For each detected fragment, quality of match (average log-likelihood, *aLL)* is plotted vs.

percentage of all classified molecules. Detection confidence threshold (total log-likelihood, *TLL* of randomized templates plus five standard deviations) is also plotted (solid line). B) *TLL* for detection of *E.coli* Sakai using sequence-derived or experimental templates, with or without elasticity correction. Error represents standard deviation of six replicate experiments. C) Classification of 1:100 mixture of *E.coli* Sakai (squares) with *Citrobacter* (circles) using SpeI digest. *E.coli* Sakai detection used linearized experimental templates while *Citrobacter* was compared against sequence-derived templates. Templates were used with (filled) or without (open) elasticity correction. *E.coli* Sakai detection above confidence threshold is emphasized in the inset. D) Percentage of expected fragments detected above confidence threshold for 1:100 mixture of *E.coli* Sakai and *Citrobacter*. Error is standard deviation of triplicate experiments.

REFERENCES

1. I. Rouzina and V. A. Bloomfield, *Biophys. J.*, 2001, 80, 894-900.
2. I. Rouzina and V. A. Bloomfield, *Biophys. J.*, 2001, 80, 882-893.
3. S. B. Smith, L. Finzi and C. Bustamante, *Science*, 1992, 258, 1122-1126.
4. T. Paramanathan, I. Vladescu, M. J. McCauley, I. Rouzina and M. C. Williams, *Nucleic Acids Res.*, 2012, 40, 4925-4932.
5. I. D. Vladescu, M. J. McCauley, I. Rouzina and M. C. Williams, *Phys. Rev. Lett.*, 2005, 95, 158102.
6. M. C. Williams, I. Rouzina, J. R. Wenner, R. J. Gorelick, K. Musier-Forsyth and V. A. Bloomfield, *Proc Natl Acad Sci U S A*, 2001, 98, 6121-6126.
7. I. D. Vladescu, M. J. McCauley, M. E. Nunez, I. Rouzina and M. C. Williams, *Nat Methods*, 2007, 4, 517-522.
8. S. K. Das, M. D. Austin, M. C. Akana, P. Deshpande, H. Cao and M. Xiao, *Nucleic Acids Res.*, 2010, 38, e177.
9. H. Yokota, F. Johnson, H. Lu, R. M. Robinson, A. M. Belu, M. D. Garrison, B. D. Ratner, B. J. Trask and D. L. Miller, *Nucleic Acids Res.*, 1997, 25, 1064-1070.
10. A. Bensimon, A. Simon, A. Chiffaudel, V. Croquette, F. Heslot and D. Bensimon, *Science*, 1994, 265, 2096-2098.
11. M. C. Williams and I. Rouzina, *Curr. Opin. Struct. Biol.*, 2002, 12, 330-336.
12. R. Dylla-Spears, J. E. Townsend, L. Jen-Jacobson, L. L. Sohn and S. J. Muller, *Lab Chip*, 2010, 10, 1543-1549.
13. J. W. Larson, G. R. Yantz, Q. Zhong, R. Charnas, C. M. D'Antoni, M. V. Gallo, K. A. Gillis, L. A. Neely, K. M. Phillips, G. G. Wong, S. R. Gullans and R. Gilmanshin, *Lab Chip*, 2006, 6, 1187-1199.

14. P. S. Doyle, B. Ladoux and J. L. Viovy, *Phys. Rev. Lett.*, 2000, 84, 4769-4772.
15. B. Ladoux and P. S. Doyle, *Europhys. Lett.*, 2000, 52, 511-517.
16. T. T. Perkins, D. E. Smith and S. Chu, *Science*, 1997, 276, 2016-2021.
17. G. C. Randall, K. M. Schultz and P. S. Doyle, *Lab Chip*, 2006, 6, 516-525.
18. N. K. Mani, S. Rudiuk and D. Baigl, *Chemical communications*, 2013, 49, 6858-6860.
19. J. W. Griffis, E. Protozanova, D. B. Cameron and R. H. Meltzer, *Lab Chip*, 2013, 13, 240-251.
20. C. U. Murade, V. Subramaniam, C. Otto and M. L. Bennink, *Biophys. J.*, 2009, 97, 835-843.
21. C. U. Murade, V. Subramaniam, C. Otto and M. L. Bennink, *Nucleic Acids Res.*, 2010, 38, 3423-3431.
22. J. A. Abels, F. Moreno-Herrero, T. van der Heijden, C. Dekker and N. H. Dekker, *Biophys. J.*, 2005, 88, 2737-2744.
23. L. S. Lerman, *Proc Natl Acad Sci U S A*, 1963, 49, 94-102.
24. R. M. Doornbos, B. G. De Grooth, Y. M. Kraan, C. J. Van Der Poel and J. Greve, *Cytometry*, 1994, 15, 267-271.
25. H. S. Rye, S. Yue, D. E. Wemmer, M. A. Quesada, R. P. Haugland, R. A. Mathies and A. N. Glazer, *Nucleic Acids Res.*, 1992, 20, 2803-2812.
26. R. Eckel, R. Ros, A. Ros, S. D. Wilking, N. Sewald and D. Anselmetti, *Biophys. J.*, 2003, 85, 1968-1973.
27. D. H. Paik and T. T. Perkins, *Angew Chem Int Ed Engl*, 2012, 51, 1811-1815.
28. K. Gunther, M. Mertig and R. Seidel, *Nucleic Acids Res.*, 2010, 38, 6526-6532.
29. A. Celedon, D. Wirtz and S. Sun, *The journal of physical chemistry. B*, 2010, 114, 16929-16935.
30. F. Johansen and J. P. Jacobsen, *J. Biomol. Struct. Dyn.*, 1998, 16, 205-222.
31. E. T. Mollova, V. A. Patil, E. Protozanova, M. Zhang and R. Gilmanshin, *Anal. Biochem.*, 2009, 391, 135-143.
32. E. Protozanova, M. Zhang, E. J. White, E. T. Mollova, D. Ten Broeck, S. V. Fridrikh, D. B. Cameron and R. Gilmanshin, *Anal. Biochem.*, 2010, 402, 83-90.
33. E. Y. Chan, N. M. Goncalves, R. A. Haeusler, A. J. Hatch, J. W. Larson, A. M. Maletta, G. R. Yantz, E. D. Carstea, M. Fuchs, G. W. Wong, S. R. Gullans and R. Gilmanshin, *Genome Res.*, 2004, 14, 1137-1146.
34. R. H. Meltzer, *The Botulinum Journal*, 2012, 2, 159-163.
35. R. H. Meltzer, J. R. Krogmeier, L. W. Kwok, R. Allen, B. Crane, J. W. Griffis, L. Knaian, N. Kojanian, G. Malkin, M. K. Nahas, V. Papkov, S. Shaikh, K. Vyavahare, Q. Zhong, Y. Zhou, J. W. Larson and R. Gilmanshin, *Lab Chip*, 2011, 11, 863-873.
36. E. J. White, S. V. Fridrikh, N. Chennagiri, D. B. Cameron, G. P. Gauvin and R. Gilmanshin, *Clin. Chem.*, 2009, 55, 2121-2129.
37. R. E. Burton, E. J. White, T. R. Foss, K. M. Phillips, R. H. Meltzer, N. Kojanian, L. W. Kwok, A. Lim, N. L. Pellerin, N. V. Mamaeva and R. Gilmanshin, *Lab Chip*, 2010, 10, 843-851.
38. A. Abibi, E. Protozanova, V. V. Demidov and M. D. Frank-Kamenetskii, *Biophys. J.*, 2004, 86, 3070-3078.
39. M. Riley, T. Abe, M. B. Arnaud, M. K. Berlyn, F. R. Blattner, R. R. Chaudhuri, J. D. Glasner, T. Horiuchi, I. M. Keseler, T. Kosuge, H. Mori, N. T. Perna, G. Plunkett, 3rd, K. E. Rudd, M. H. Serres, G. H. Thomas, N. R. Thomson, D. Wishart and B. L. Wanner, *Nucleic Acids Res.*, 2006, 34, 1-9.
40. F. R. Blattner, G. Plunkett, 3rd, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau and Y. Shao, *Science*, 1997, 277, 1453-1462.

41. M. McClelland, K. E. Sanderson, J. Spieth, S. W. Clifton, P. Latreille, L. Courtney, S. Porwollik, J. Ali, M. Dante, F. Du, S. Hou, D. Layman, S. Leonard, C. Nguyen, K. Scott, A. Holmes, N. Grewal, E. Mulvaney, E. Ryan, H. Sun, L. Florea, W. Miller, T. Stoneking, M. Nhan, R. Waterston and R. K. Wilson, *Nature*, 2001, 413, 852-856.
42. M. D. Abramoff, P. J. Magalhaes and S. J. Ram, *Biophotonics International*, 2004, 11, 36-42.
43. J. L. Seifert, R. E. Connor, S. A. Kushon, M. Wang and e. B. A. Armitag, *J. Am. Chem. Soc*, 1999, 121, 2987-2995.
44. M. Doi and S. F. Edwards, *The Theory of Polymer Dynamics*, Oxford University Press, Oxford, 1986.
45. J. D. Watson and F. H. Crick, *Nature*, 1953, 171, 737-738.
46. A. N. Glazer and H. S. Rye, *Nature*, 1992, 359, 859-861.
47. T. Anazawa, H. Matsunaga and E. S. Yeung, *Anal. Chem.*, 2002, 74, 5033-5038.
48. H. Hoshi, J. O'Brien and S. L. Mills, *J. Histochem. Cytochem.*, 2006, 54, 1169-1176.
49. T. Suzuki, K. Fujikura, T. Higashiyama and K. Takata, *J. Histochem. Cytochem.*, 1997, 45, 49-53.
50. H. Gudnason, M. Dufva, D. D. Bang and A. Wolff, *Nucleic Acids Res.*, 2007, 35, e127.
51. K. D. Dorfman, *AIChE J.*, 2013, 59, 346-354.
52. X. Zhang, H. Chen, H. Fu, P. S. Doyle and J. Yan, *Proc Natl Acad Sci U S A*, 2012, 109, 8103-8108.
53. S. Winter and G. Loeber, *Journal of biomedical optics*, 1997, 2, 125-130.

Figure 1

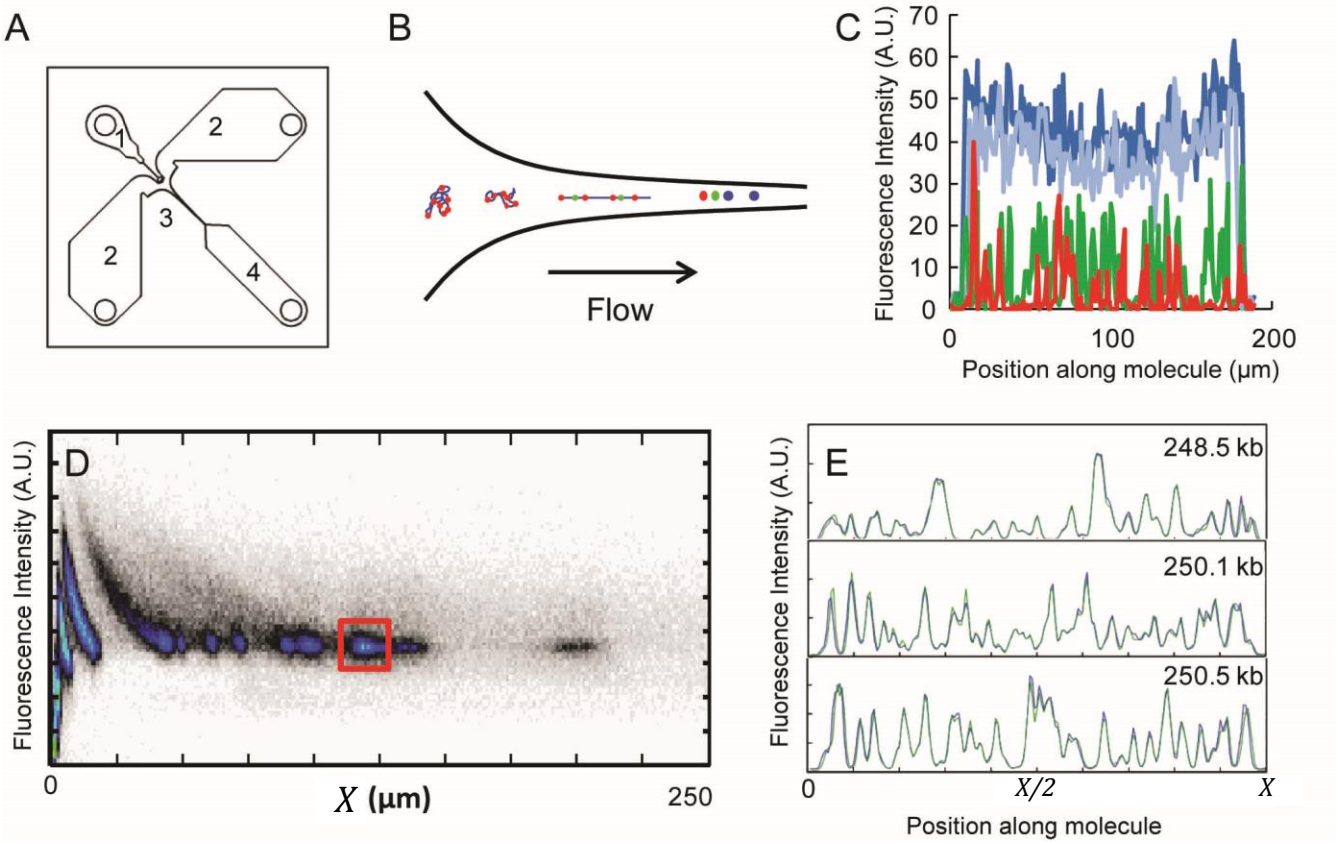

Figure 2

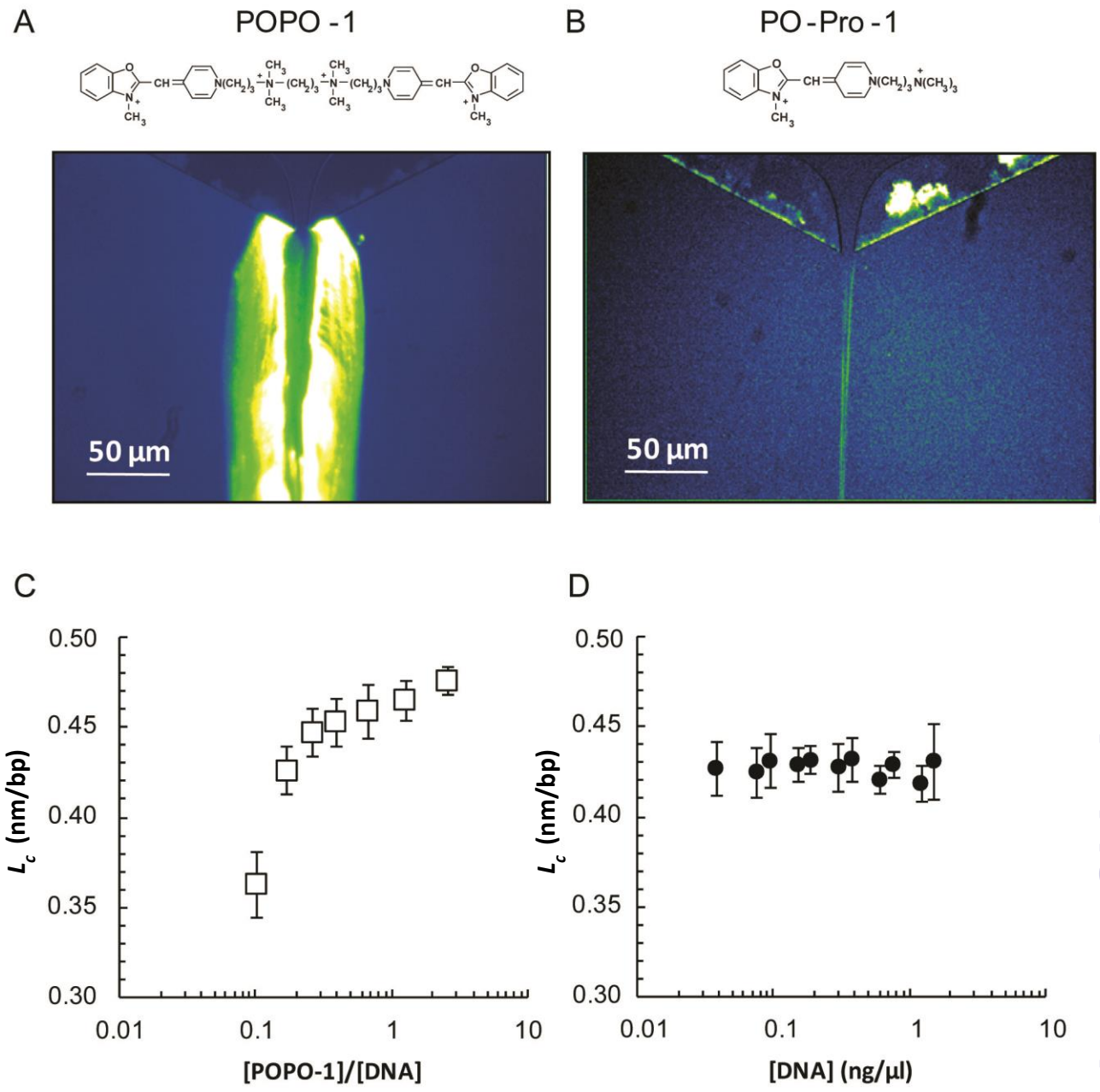

Figure 3

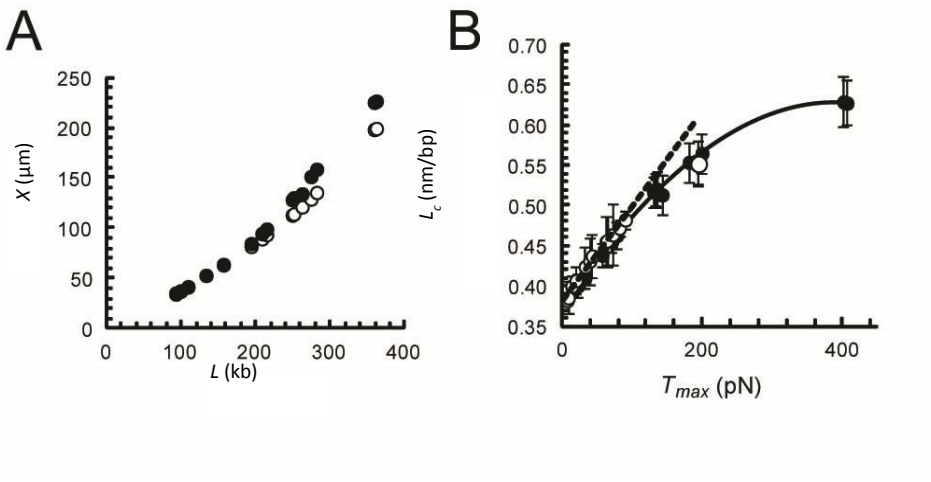

Figure 4

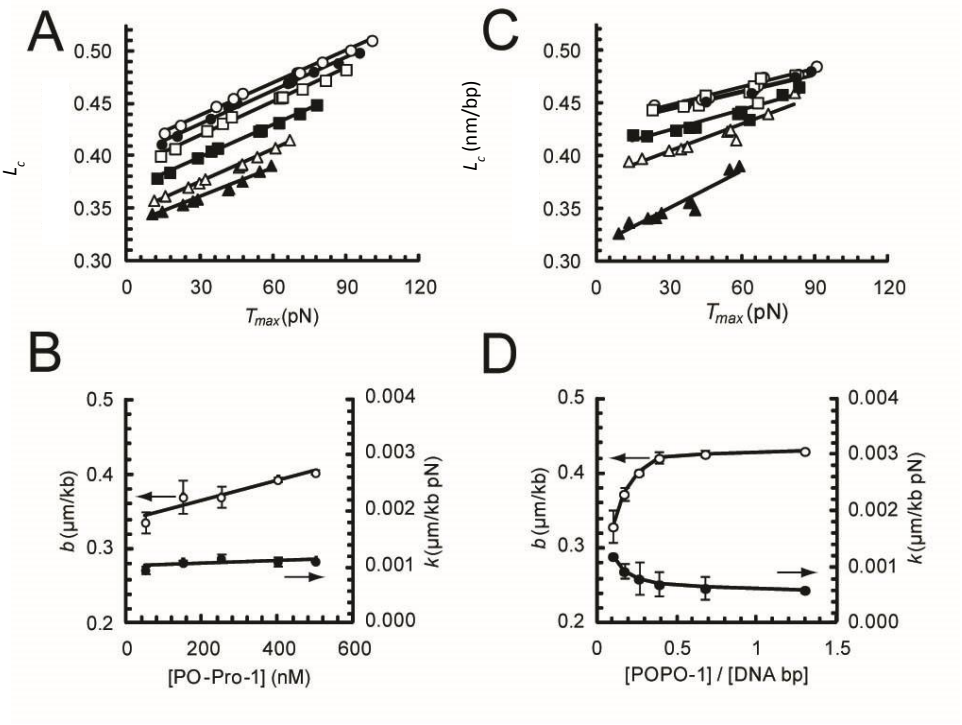

Figure 5

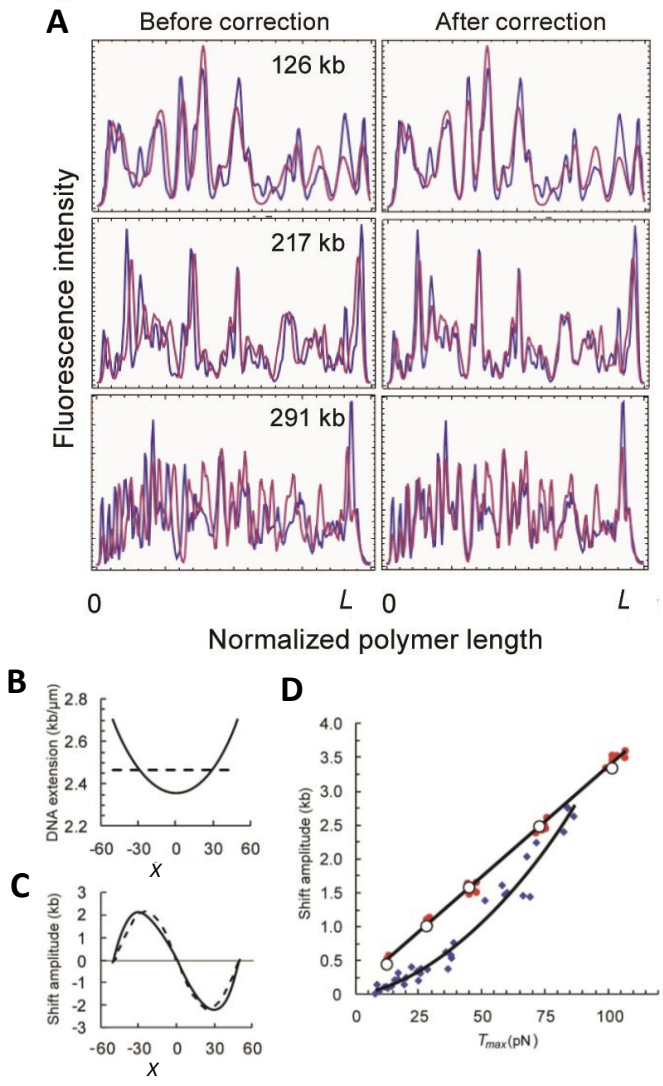

Figure 6