Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Chromatography A, 1053 (2004) 173–179
JOURNAL OF CHROMATOGRAPHY A
www.elsevier.com/locate/chroma

Single-molecule spectroscopy for molecular identification in capillary electrophoresis

Ji-young Lee, Hung-Wing Li, Edward S. Yeung*

Ames Laboratory—US Department of Energy and Department of Chemistry, Iowa State University, Ames, IA 50011, USA

Available online 28 July 2004

Abstract

The electrophoretic mobilities of individual DNA molecules were determined by using fluorescence imaging. The average mobility agreed well with that obtained in capillary electrophoresis (CE). The signal-to-noise ratio (S/N) did not decrease in the presence of up to 8% plasma or 8% raw blood. Single-molecule detection was still possible in the presence of 50% raw blood. Single-molecule CE of two differently labeled molecules was carried out in the presence of a transmission grating. Even when the mobility difference is not sufficient because of low S/N, identification using different fluorescence wavelengths can be performed at >99% accuracy. So, when small differences in DNA sequence due to disease or mutation can lead to hybridization to labels with different dyes, the screening of the mutated DNA will be facilitated by online spectroscopy in addition to the electrophoretic information from CE.

Keywords: Single-molecule detection; DNA; Dyes

1. Introduction

The need for ultrasensitive detection techniques in clinical research, down to the single-molecule limit, is best illustrated by the recent findings in acquired immunodeficiency syndrome (AIDS) research. Previously, there was excitement in the medical community that a combination of two drugs that attack the reverse transcriptase enzyme and one drug that attacks the protease enzyme can drive the viral load in human immunodeficiency virus (HIV) infected individuals down to a negligible level [1]. At the time, the threshold for detection is 200–400 copies of the virus per mL of plasma [1]. Recently, the detectable level of HIV in plasma has been improved to as low as 20 copies/mL [2]. In a 151-person study, samples from the 27% of patients who originally had undetectable levels (<400 copies/mL) were reanalyzed. Only 12% had unquantifiable levels with the new test (<20 copies/mL). Furthermore, for one of the samples below 20 copies/mL, the virus was detected three out of 11 times [3]. This shows how important it is to develop even more sensitive tests for infection. Ultimately, detection of HIV down to the single copy level will be desirable. Since progression of HIV infection to AIDS, prognosis, and treatment have often been tied to viral load [4,5], such ultrasensitive detection technology is badly needed for early detection. Even if 20 copies are present per mL, one would like to know the diversity among them, since mutations are common. Single-molecule assays offer such information.

Our approach [6–10] is based on fluorescence imaging of individual molecules within a thin solution film. In addition, it will be necessary to selectively image the targeted species in an overwhelming excess of biogenic or probe molecules. To become practical in the clinical laboratory, one must have the required sensitivity, have 100% counting efficiency, and be able to do so at a high rate.

Many assays relevant to disease diagnosis are based on electrophoresis, where the migration velocity is used for distinguishing molecules of different size or charge. However, standard gel electrophoresis is not only slow but also insensitive. The single-molecule imaging procedure can measure the electrophoretic mobilities of many distinct molecules every second. The molecules are not separated as in conventional electrophoresis. They are simply identified and counted with

* Corresponding author. Tel.: +1 515 2946 342; fax: +1 515 2940 105.
E-mail address: yeung@ameslab.gov (E.S. Yeung).

0021-9673/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2004.06.118

close to 100% efficiency. This means that virtually all electrophoresis diagnostic protocols from slab gels to capillary electrophoresis should be adaptable to single-molecule detection.

Recently, we developed a high-speed high-throughput single-molecule imaging technique for identifying molecules that are labeled with different dyes in free solution based on differences in their fluorescence emission spectra [11]. Unlike previous reports, the entire spectrum, rather than selected wavelengths through optical filters, is recorded. Our system employs a transmission grating for wavelength dispersion. The complete emission spectra of many individual molecules are imaged in the ms time scale and recorded simultaneously even though they are freely moving in the observation region. This technique can therefore be used for screening single molecules for disease markers and for monitoring individual molecular interactions at a rate of thousands of molecules per second.

In this work, we demonstrate that high throughput measurement of single-molecule electrophoretic mobilities is possible even in the presence of raw blood or plasma. We further combine the best features of electrophoretic and spectroscopic discrimination among single molecules. At high signal-to-noise (S/N) conditions, electrophoresis provides confident identification of the molecules. At low S/N, the spectroscopic information allows identification even when the electrophoretic mobilities are similar.

2. Experimental

2.1. Buffer solutions

C16E6 is a monomeric non-ionic surfactant used as the sieving matrix for the separation of DNA fragments by capillary electrophoresis [12]. A 0.25% (w/w) C16E6 (Sigma, St. Louis, MO, USA) prepared in 10 mM Gly–Gly buffer (Sigma) was used in all experiments. Gly–Gly buffer solution was dissolved in ultrapure water and adjusted to pH 8.2 by dropwise addition of 2 M NaOH. The buffer solution was filtered with 0.2 μm filter. C16E6 was dissolved in Gly–Gly buffer at 90 °C and cooled down to room temperature with constant stirring at medium speed. The C16E6 solution was also filtered with 0.2 μm filter before use.

2.2. DNA Samples

Lambda-DNA [48502 base pairs (bp), New England Biolabs, MA, USA] was labeled with YOYO-1 (quinolinium, 1,1′-{1,3-propanediylbis[(dimethyliminio)-3,1-propanediyl]}bis{4-[(3-methyl-2(3H)-benzoxazolylidene)methyl]}-, tetraiodide) or POPO-III (benzoxazolium, 2,2′[1,3-propanediylbis[dimethyliminio)-3,1-propanediyl-1(4H)-pyridinyl-4-ylidene-1-propen-1-yl-3-ylidene]]bis[3-methyl]-tetraiodide, Molecular Probes, Eugene, OR, USA) at a ratio of one dye molecule per five base pairs according to the manufacturer's instructions. Labeled Lambda DNA samples were first prepared as a 500 pM stock solution in 10 mM Gly–Gly (Sigma) buffer, pH 8.2, incubated at room temperature for at least 1 h, and diluted with ultrapure (18 MΩ) water to 0.5, 0.2 and 0.1 pM just prior to the start of the experiment.

For electrophoresis and spectroscopic experiments, M13mp18 (7249 bp) and M13KE (7222 bp) DNA were obtained from New England Biolabs. Both DNAs were cut with Acc65I and BsmI restriction enzymes to result in 4497 and 2752 bp DNA fragments for M13mp18 and 7088 bp and 135 bp for M13KE. The resulting fragments were also labeled fluorescently; M13mp18 fragments were labeled with YOYO-I and M13KE fragments with POPO-III. Both YOYO-I and POPO-III dyes were used at a ratio of one dye molecule per five base pairs according to the manufacturer's instructions. All samples were allowed to incubate for 5 min before they were further diluted. Lambda DNA were prepared at a concentration of 200 pM and diluted to 0.2 pM. M13mp18 and M13KE DNA samples were prepared at 1.4 nM and diluted to 2.8 pM in 0.25% (w/w) C16E6 solution prior to experiments.

2.3. Raw blood and blood plasma solutions

Raw blood and blood plasma obtained from healthy adult donors were first diluted to 50 and 10% of the original concentration with ultrapure water. Table 1 shows the volumes of the stock solutions that were used to prepare the sample solutions. These volumes were chosen so that the final concentrations of DNA in each sample solution were identical.

2.4. Instrument setup

A 27 cm long square capillary (75 μm i.d. × 365 μm o.d.; Polymicro Technologies, Phoenix, AZ, USA) was used for all electrophoresis experiments. The capillary was mounted on a bakalite block, which was placed on a Zeiss Axioskop upright microscope. The light source and detection system were very similar to those used in previous studies done in our group [10,11,13,14]. A Uniphase (San Jose, CA, USA) argon ion laser operating at 488 nm with 12.2 mW output power was used as the excitation source. The laser beam was focused by a 2.5 cm focal length cylindrical lens (Edmund Industrial Optics, Barrington, NJ, USA) so that the focal point was at the center of the capillary. Fluorescence from single DNA molecules was collected with a Zeiss 20× Fluar (0.75 NA) microscope objective at right angle to the incident laser beam. Two 488 nm holographic notch filters (Keiser Optical, Ann Arbor, MI, USA; HNFP) with optical density >6 were placed between the microscope objective and a Pentamax 512-EFT/1EIA intensified charge-coupled device camera (ICCD, Princeton Instruments, Princeton, NJ, USA). The ICCD collected images with 30 ms exposure time at 2 Hz. The laser beam was chopped synchronously to the

Table 1
Preparation of various sample solutions for single-molecule DNA imaging

Solution	YOYO-I-labeled λ-DNA (pM)			Blood plasma (%)			Raw blood (%)	
	0.1	0.2	0.5	10	50	100	10	100
1	1000 μL							
2			200 μL	800 μL				
3		500 μL				500 μL		
4	500 μL				500 μL			
5			200 μL				800 μL	
6		500 μL						500 μL

ICCD image collection rate with a Uniblitz mechanical shutter (LS2Z2, Vincent Associates, Rochester, NY, USA) and a driver (model T132, Vincent Associates) to reduce photobleaching. A transmission grating with 70 grooves/mm (Edmund) was mounted in front of the ICCD camera to disperse fluorescence light from the molecules. The distance between the transmission grating and the ICCD was set to be 4 cm so that the zeroth-order and first-order images of the 75 μm capillary could not overlap each other [11].

2.5. Electrophoresis conditions

The capillary was equilibrated with 0.25% C16E6 in 50 mM Gly–Gly solution (pH 8.2) for 15 min before use. The entire capillary was filled with sample solution with a syringe 5 min before irradiation by the excitation beam. Before each electrophoresis experiment, the heights of the buffer reservoirs were adjusted while molecular movement was monitored until no hydrodynamic flow existed in the capillary. Electric field, 1 kV/27.0 cm d.c., was applied with a Spellman high-voltage d.c. supply (Hauppauge, NY, USA). Data collection started 10 s after applying voltage. Between electrophoresis experiments, the capillary was washed with 0.25% C16E6 in Gly–Gly solution for 5 min.

3. Results and discussion

3.1. Image analysis

The electrophoresis experiment in Fig. 1 shows that the single-molecule mobility can be used for identification even though the individual intensities are very different and are not constant with time. This is a clear advantage over intensity-based measurements [15]. In Fig. 1, as many as 50 × 50 molecules can fit into the full 512 × 512 CCD frame. Therefore, 2,500 molecules can be screened in 25 ms. The observation volume is about 10^{-2} μL. Another factor of a 10–50 increase in throughput is expected, if larger imaged areas, higher electric fields, faster frame rates and higher laser powers are employed. The use of a wide but thin flow channel coupled with a wide-field microscope objective and synchronization of the applied voltage with the frame rate will

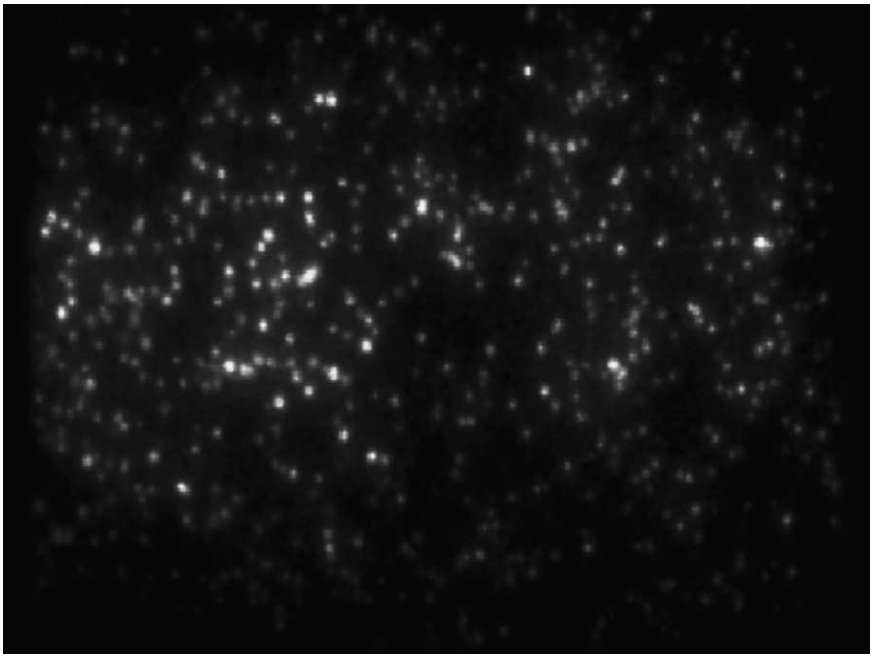

Fig. 1. High-throughput imaging of single YOYO-I-labeled DNA molecules. Each bright spot above the dark background corresponds to an individual molecule.

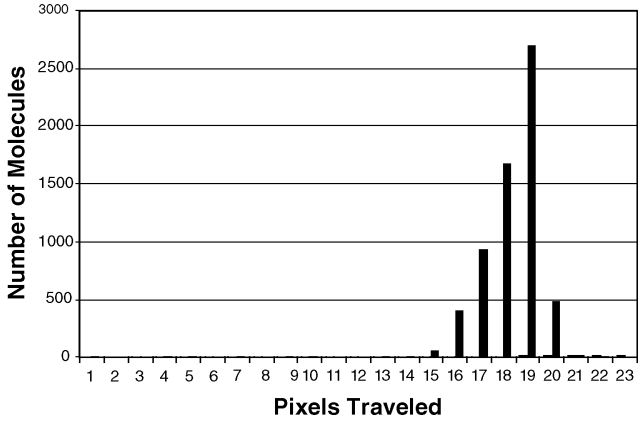

Fig. 2. Histogram of measured migration distances between any two frames in the series of images from Fig. 1.

guarantee that 100% of the molecules are imaged. Therefore, one should in principle be able to screen one drop of blood (40 μL) in 10 s.

In this study, the consecutive images were examined automatically and the movements were tracked by counting the number of pixels traveled by each molecule over a known time interval (frame number). Molecules were counted when their intensities exceeded a selected threshold value above the background intensity. We simply cross-correlated the sequences of images by multiplying them with each other after applying well-defined shifts in pixel number. Briefly, the pixels in each image in the movies were first converted into either 1 or 0 depending on whether the intensity was above or below a selected threshold. This resulted in a one-bit image that facilitates computation. The entire image in the sequence was then shifted forward in row numbers (electrophoresis direction) by x pixels. The resulting shifted image was multiplied by the $(n + 1)$ image in the sequence, pixel by pixel. The rare event (bound complex or deleted DNA fragment) with the prescribed electrophoretic mobility will then be highlighted in the resulting cross-correlation image (1×1) while the normal events in large excess (probe molecule or intact DNA) will be suppressed (1×0 or 0×1). Data analysis can, thus, be performed at rates comparable to data acquisition. The results of such a computer-generated cross-correlation are shown in Fig. 2. The calculated mobility for these molecules is $1.08 \times 10^{-4}\,\mathrm{cm^2/V\,s^{-1}}$, exactly as predicted from bulk electrophoresis experiments. The relative standard deviation is 7.1%.

Images from each sample solution were evaluated by WinView software (Roper Scientific, Princeton, NJ, USA). The typical maximum intensity of the molecule spots constituted the signal (s). The average (b) and standard deviation (dev) of the background were determined from molecule-free areas in the same image. The results are shown in Table 2. The last column there is equivalent to the traditional signal-to-noise (S/N) ratio. We can see that S/N was actually higher when 8% plasma was present (solution 2). There was an unexpected increase in the signal level (almost 4×), which more than compensated for the increase in the background fluorescence from the other biological materials. This is likely due to the stabilization of the fluorophor by proteins that adsorb onto the DNA. The weak adsorption apparently did not change the mobility of the DNA because of its high negative charge and large size. The background level further increased when raw blood was present instead of plasma, but at 8% raw blood (solution 5) the net S/N is still comparable to DNA molecules in pure buffer (solution 1).

The S/N for DNA imaging in both plasma and raw blood did deteriorate at high fractions of either. However, even at 50%, the S/N of each was still above the required value of three that defines the detection limit. Molecule counting at low S/N is demonstrated in Fig. 3. The three images in Fig. 3a were plotted with the contrast levels indicated in Table 1 for the signal and the background, respectively. The images clearly show that S/N is best in 8% plasma, while 8% raw blood provided identical S/N to a pure buffer. After readjusting the threshold levels and the signal levels under each condition for confident identification of molecules, the three images in Fig. 3b were obtained. These contrast levels parallel the ones selected automatically by standard chromatography software for peak picking. Roughly the same number of DNA molecules were found in each image. This confirms that our protocol is suitable for imaging biological fluids with little sample preparation or dilution.

3.2. Fluorescence imaging with transmission grating

A fluorescence image from single DNA molecules in a capillary dispersed with the transmission grating is shown in Fig. 4. The capillary was filled with 1:1 mixture of YOYO-I-labeled Lambda DNA and POPO-III-labeled Lambda DNA, both dissolved in 0.25% (w/w) C16E6 in 50 mM Gly–Gly buffer (pH 8.2). Since POPO-III has an absorption maximum at 534 nm and has an efficiency of 26% at 488 nm, POPO-III-

Table 2
Comparison of DNA signal-to-noise ratios in various solutions described in Table 1

Solution	Signal counts (s)	Background counts (bg)	Noise deviation counts (dev)	(s-bg)/bg	(s-bg)/dev
1	1100	250	35	3.40	24.3
2	4000	700	60	4.71	55.0
3	3500	2300	170	0.52	7.1
4	3700	1900	150	0.95	12.0
5	3300	900	100	2.67	24.0
6	4000	2800	300	0.43	4.0

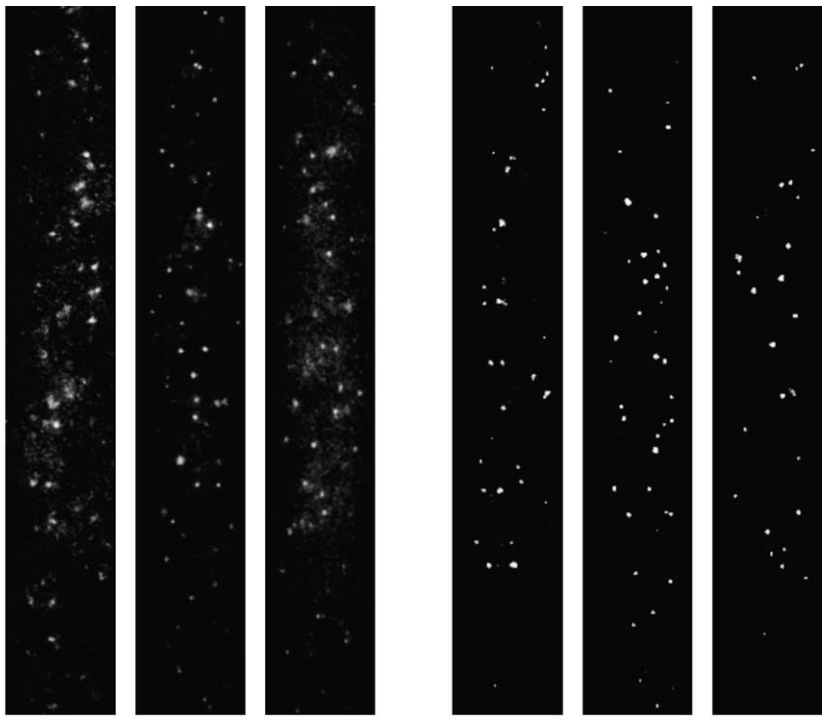

Fig. 3. Discrimination of single DNA molecules from background fluorescence: (a) raw data of molecules on a fluorescent background and; (b) normalized data allowing detection of molecules without background interference.

labeled DNAs were not as bright as YOYO-I-labeled DNAs that have absorption maximum at 491 nm. With the distance of 4 cm between the transmission grating and the ICCD, the non-dispersed zeroth-order image (left half) of the 75 μm capillary is imaged simultaneously with the dispersed first-order image (right half).

The zeroth-order image is identical to that without transmission grating and can be used to calculate the mobility of individual DNA in the sieving matrix under electric field, just as in capillary gel electrophoresis. The zeroth-order capillary image was also used as a reference point to determine the dispersion distance between zeroth-order and first-order images. The emission maximum at 509 nm for YOYO-I and 570 nm for POPO-III resulted in dispersion distances of 1.425 mm for YOYO-I and 1.597 mm for POPO-III at this seting. Counting the number of pixels between the highest intensity pixels in the zeroth-order and corresponding first-order images allows the molecules to be distinguished from each other.

3.3. DNA fragments after restriction enzyme reaction

M13KE is a derivative of M13mp19 and is different from M13mp18 by 26 bp. These two DNAs mimic a situation with normal DNA and a DNA with mutation caused by disease. This small difference, however, creates different cutting sites for Acc65 I enzyme. BsmI cuts both at the same location to open the circular DNAs. As a result, M13KE cleaves into 7088 and 135 bp fragments, while M13mp18 DNA gives 4497 and 2752 bp fragments after cutting with Acc65I and BsmI restriction enzymes.

3.4. Electrophoretic and spectroscopic data

Fig. 5 is a histogram of mobility of POPO-III-labeled M13KE and YOYO-I-labeled M13mp18 when the laser power is not high enough to provide adequate S/N ratio. The low S/N condition provides a stringent test of our

Table 3
Discrimination between DNA by electrophoresis and by spectroscopy

	Electrophoretic mobility		Wavelength dispersion	
	YOYO-I-labeled M13mp18 (pixels)	POPO-III-labeled M13KE (pixels)	YOYO-I-labeled M13mp18 (pixels)	POPO-III-labeled M13KE (pixels)
Average	60.4	46.6	79.5	87.9
Standard deviation	8.58	21.3	1.78	2.16

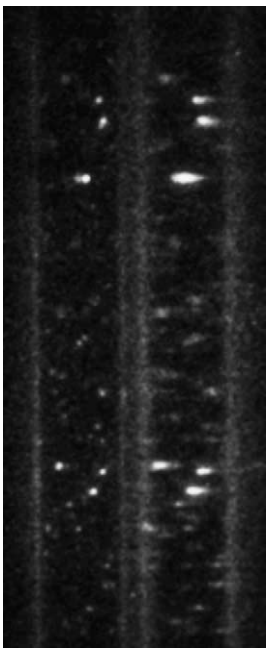

Fig. 4. Fluorescence emission from single molecules dispersed with a 70 grooves/mm transmission grating. 1:1 Mixture of YOYO-I-labeled Lambda DNA and POPO-III-labeled Lambda DNA were filled in a 75-μm square capillary. Left, undispersed zeroth-order image. Each spot corresponds to one molecule. Right, dispersed first-order image spaced from the zero-order according to the wavelength. Each horizontal streak corresponds to a dispersed spectrum.

discrimination scheme. The low laser power is required to avoid photobleaching over the series of successive images. The *x*-axis is the number of pixels traveled by the molecules in 1 s and the *y*-axis is the number of molecules. Picking

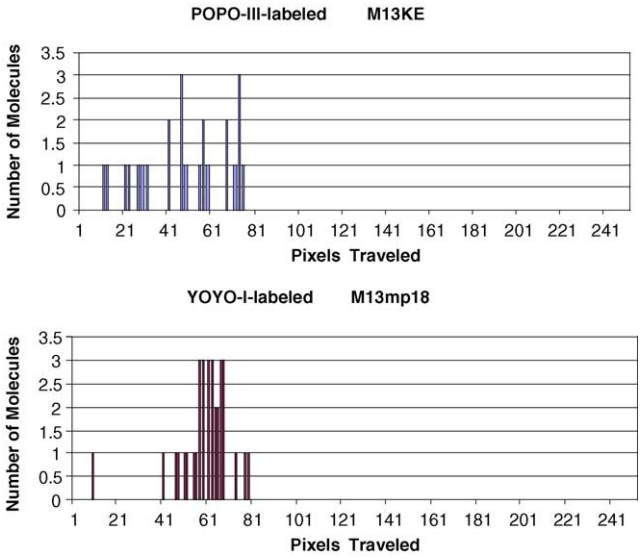

Fig. 5. Histograms of DNA mobilities for POPO-III-labeled M13KE restriction enzyme digest and YOYO-I-labeled M13mp18 restriction enzyme digest.

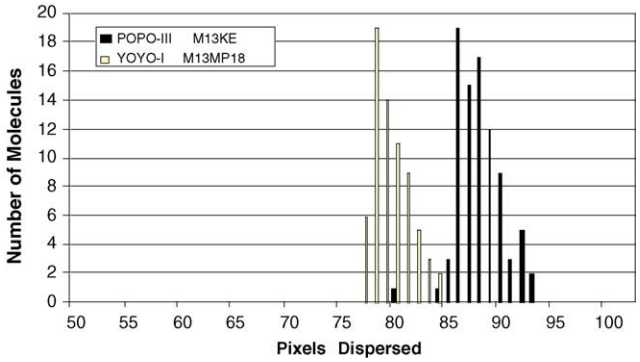

Fig. 6. Distribution of fluorescence maxima from POPO-III-labeled M13KE (7088 bp) and YOYO-I-labeled M13mp18 (4497 and 2752 bp) fragments.

molecules in these movies and calculation of their mobilities were done with the in-house software described in the image analysis section. Only the zeroth-order image area was subjected to the calculation of mobilities. Since the smallest DNA fragment that can be seen in this setup is ∼800 bp [9], the 135 bp fragment from M13KE is not shown. As shown in Fig. 5, the mobility differences found for 7088 bp M13KE and 4497 bp and 2752 bp M13mp18 were not distinct. With S/N = 2, the 2752 bp fragments were not picked up by the software. Table 3 shows that lower S/N from the less efficient emitter, POPO-III-labeled M13KE, resulted in a broad range of measured mobilities (with a standard deviation of almost 50%), which even covers the entire M13mp18 mobility region. As is, it is not possible to identify molecules simply by mobility.

Fig. 6 is the wavelength dispersion distribution of POPO-III-labeled M13KE and YOYO-I-labeled M13mp18. Higher laser powers can be used since only one image is required for spectral determination. From Table 3, the average dispersion of POPO-III is 87.9 pixels with a standard deviation of 2.16 and that of YOYO-I is 79.5 pixels with a standard deviation of 1.78. The difference between the two averages is 8.43, which is 4.3 times the average of the two standard deviations. With more than four times the standard deviation, the two peaks are >99% separated from each other [16]. In other words, when DNA molecules are labeled with different dye colors, their identification can be done with greater than 99% accuracy.

Acknowledgments

We thank Gang Xue for developing the in-house image analysis software. E.S.Y. thanks the Robert Allen Wright Endowment for Excellence for support. The Ames Laboratory is operated for the US Department of Energy by Iowa State University under Contract No. W-7405-Eng-82. This work was supported by the Director of Science, Office of Basic Energy Sciences, Division of Chemical Sciences and by the National Institutes of Health.

References

[1] D.D. Richman, Science 272 (1996) 1886.
[2] J. Cohen, Science 276 (1997) 1488.
[3] J. Cohen, Science 277 (1997) 32.
[4] D.D. Ho, Science 272 (1996) 1124.
[5] J.W. Mellors, J. Rinaldo, R. Charles, P. Gupta, R.M. White, J.A. Todd, L.A. Kingsley, Science 272 (1996) 1167.
[6] M.R. Shortreed, H. Li, W.-H. Huang, E.S. Yeung, Anal. Chem. 72 (2000) 2879.
[7] M.R. Shortreed, H. Li, E.S. Yeung, Scanning and Force Microscopies For Biomedical Applications II, in: S. Nie, E. Tamiya, E.S. Yeung (Eds.), Proceedings of SPIE, Vol. 3922, 2000, p. 58.
[8] Y. Ma, M.R. Shortreed, H. Li, W. Huang, E.S. Yeung, Electrophoresis 22 (2001) 421.
[9] T. Anazawa, H. Matsunaga, E.S. Yeung, Anal. Chem. 74 (2002) 5033.
[10] J. Zheng, E.S. Yeung, Aust. J. Chem. 56 (2003) 149.
[11] Y. Ma, M.R. Shortreed, E.S. Yeung, Anal. Chem. 72 (2000) 4640.
[12] W. Wei, E.S. Yeung, Anal. Chem. 73 (2001) 1776.
[13] J. Zheng, E.S. Yeung, Anal. Chem. 75 (2003) 3675.
[14] J. Zheng, E.S. Yeung, Anal. Chem. 74 (2002) 4536.
[15] A. Van Orden, R.A. Keller, W.P. Ambrose, Anal. Chem. 72 (2000) 37.
[16] P.R. Bevington, D.K. Robinson, Data Reduction and Error Analysis for the Physical Sciences, McGraw-Hill, 1992.