Anal. Chem. **2000,** *72,* 5865−5873

Articles

DNA Sequencing by Capillary Electrophoresis with Four-Decay Fluorescence Detection

Hui He and Linda B. McGown*

Department of Chemistry, P. M. Gross Chemical Laboratory, Box 90346, Duke University, Durham, North Carolina 27708

A scheme for multiplex detection of dye-labeled DNA fragments in DNA sequencing is described in which on-the-fly, frequency-domain fluorescence lifetime detection is used to discriminate among the dye-labeled fragments of the four terminal bases in a single-lane CE separation. Two four-dye systems were evaluated, one excited at 488 nm and the other, at 514 nm. The 488 nm system proved successful for four-decay detection. Base calling was achieved either directly from on-the-fly lifetimes or from lifetime-resolved electropherograms recovered for each base from the electropherogram of the mixture of sequencing reaction products. The latter method was found to be more accurate (99% for two bases and 98.5% for three bases) and could achieve longer read lengths, but it was unsuccessful for sequencing of all four bases. The first method gave a base-calling accuracy of 96% for four-base sequencing over the fragment length range of 41−220 bases.

Fluorescence detection for DNA sequencing has achieved widespread use since its introduction in 1986.[1] Fluorescent labels are a relatively safe alternative to radioactive labels, providing similar detection limits, high sensitivity, ease of automation, and on-line detection.[2,3] Moreover, the selectivity of fluorescence detection makes it suitable for multiplex detection, that is, determination of multiple analytes in a single run in capillary electrophoresis (CE). Multiplex methods for DNA sequencing detection not only improve the throughput of sequencing by a factor of 4 but also avoid the problem of lane-to-lane misalignment.[4]

The first multiplex, fluorescence-based sequencing system used a four-color detection scheme with a rotating filter wheel.[1] The dyes included fluorescein isothiocyanate, NBD-aminohexanoic acid, tetramethylrhodamine isothiocyanate, and Texas Red. Other multiplex methods include two-channel detection using a set of four succinylfluorescein dyes in which the ratio of the fluorescence intensity in the two channels is used to identify the base;[5] single color, intensity-based detection using a single fluorescein dye;[6,7] four-color, four-channel detection in which the emission signal is selected through four different narrow-band-pass filters coupled with four PMTs;[8] and four-color detection using a spectrograph for spectral discrimination coupled with CCD detection.[9] Energy transfer has been used to maximize excitation efficiency using one or more laser lines for excitation.[10−13] A donor dye, FAM, is excited at 488 nm, and emission from four acceptor dyes provides color-specific identification of the bases.

In relying on color discrimination in four-color schemes, it is necessary to achieve wavelength selectivity in the fluorescence emission spectrum using one or more excitation wavelengths. Detection using only filter selection of color may be inadequate

* To whom correspondence should be addressed. Phone: (919) 660-1545. Fax: (919) 660-1605. E-mail: lbmcgown@chem.duke.edu.

(1) Smith, L. M.; Sander, J. Z.; Kaiser, R. J.; Hughes, P.; Dodd, C.; Connell, C. R.; Heiner, C.; Kent, S. B. H.; Hood, L. E. *Nature* **1986**, *321*, 674.
(2) Smith, L. M.; Fung, S.; Hunkapiller, M. W.; Hunkapiller, T. J.; Hood, L. E. *Nucleic Acids Res.* **1986**, *13*, 2399.
(3) Connell, C.; Fung, S.; Heiner, C.; Bridgham, J.; Chakerian, V.; Heron, E.; Jones, B.; Menchen, S.; Mordan, W.; Raff, M.; Recknor, M.; Smith, L. M.; Springer, J.; Woo, S.; Hunkapiller, M. W. *BioTechniques* **1987**, *5*, 342.
(4) Smith, L. M. *Genome* **1989**, *31*, 929.
(5) Prober, J. M.; Trainor, G. L.; Dam, R. L.; Hobbsk, F. W.; Robertson, C. W.; Zagursky, R. J.; Cocuzza, A. J.; Jensen, M. A.; Baumeister, K. *Science* **1987**, *238*, 336.
(6) Tabor, S.; Richardson, C. C.; *J. Biol. Chem.* **1990**, *265*, 8322.
(7) Ansorge, W.; Zimmermann, J.; Schwager, C.; Stegemann, J.; Erfle, H.; Voss, H. *Nucleic Acids Res.* **1990**, *18*, 3419.
(8) Luckey, J. A.; Drossman, H.; Kostichka, A. J.; Mead, D. A.; D'Cunha, J.; Norris, T. B.; Smith, L. M. *Nucleic Acids Res.* **1990**, *18*, 4417.
(9) Karger, A. E.; Harris, J. M.; Gesteland, R. F. *Nucleic Acids Res.* **1991**, *19*, 4955.
(10) Ju, J.; Ruan, C.; Fuller, C. W.; Glazer, A. N.; Mathies, R. A.; *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 4347.
(11) Ju, J.; Glazer, A. N.; Mathies, R. A. *Nat. Med.* **1996**, *2*, 246.
(12) Lee, L. G.; Spergeon, S. L.; Heiner, C. R.; Benson, S. C.; Rosenblum, B. B.; Menchen, S. M.; Graham, R. J.; Constantinescu, A.; Upadhya, K. G.; Cassel, J. M. *Nucleic Acids Res.* **1997**, *25*, 2816.
(13) Rosenblum, B. B.; Lee, L. G.; Spurgeon, S. L.; Khan, S. H.; Menchen, S. M.; Heiner, C. R.; Chen, S. M. *Nucleic Acids Res.* **1997**, *25*, 4500.

10.1021/ac000952n CCC: $19.00
Published on Web 11/16/2000

to resolve overlapping peaks, while the use of array detection to collect a spectrum in order to better identify and resolve colors can decrease sensitivity because of dispersion. Moreover, because four-color sequencing relies on spectral resolution of four dyes, broadly overlapping emission spectra may result in cross-talk between detection channels. Sets of four dyes with narrow emission spectra have been used to increase spectral resolution,[12,14] but some overlap is inevitable due to the continuous nature of the spectra.

Thus, it has been desirable to explore new approaches to fluorescence detection that have the potential to increase accuracy and resolution in DNA sequencing. One such approach is to use fluorescence lifetime instead of color to discriminate among the four bases. Fluorescence lifetime is a discrete function, generally governed by simple, first-order kinetics for each decay component, and is, therefore, easier to model and resolve than spectral color.[15]

The use of fluorescence lifetime detection has been discussed in the development of near-infrared dyes and time-domain detection strategies.[16−21] Dye-labeled fragments were identified by either pattern recognition of the decay profile or lifetime recovered from a monoexponential maximum likelihood estimator. Because a decay pattern in on-the-fly, time-domain lifetime detection is constructed from only a small number of photon counts, it is difficult to extract multiple lifetimes from the decay pattern, even when the actual decay is more complex.[17−22] This eliminates the advantage of on-the-fly lifetime detection for resolution of overlapping peaks of different dye-labeled fragments in order to achieve base-calling in regions of peak overlap.

An alternative to time-domain detection in capillary electrophoresis (CE) has been described in which fluorescence lifetime is measured in the frequency domain using commercial, multiharmonic Fourier transform phase-modulation instrumentation.[15,23−25] This instrumental approach has been shown to provide unprecedented speed for on-the-fly detection and subsequent lifetime resolution of overlapping peaks in CE. An entire frequency response, which contains sufficient information to analyze multiexponential decays, can be acquired on the order of milliseconds, providing many points per electrophoretic peak. We have resolved as many as three overlapping peaks in CE separations of dye-labeled DNA primers by using on-the-fly lifetime detection.[25]

We describe here the use of fluorescence lifetime detection in CE for DNA sequencing in an approach that has been referred to as "four-decay" detection.[15,24,25] Previously, several sets of four dye-labeled primers were characterized for application in four-decay sequencing,[24,25] but they did not prove successful in actual sequencing applications. In the present work, two new four-dye systems were investigated. One is excited at 488 nm and consists of Cy3, fluorescein−dTMR, rhodamine green (RG), and BODIPY-FL (BOD). The other is excited at 514 nm and consists of Cy3, BODIPY-FL Br_2 (BBr), tetramethyl rhodamine (TMR), and BODIPY Rhodamine 6G (BR6G). Two base-calling methods for DNA sequencing were compared. In one, the lifetimes recovered for each peak were used directly to identify the peak. In the other, lifetime-resolved electropherograms were recovered for each base from the electropherogram of the mixture of sequencing reaction products.

EXPERIMENTAL SECTION

Reagents. Dye-labeled M13-forward primers d(5′-TGT-AAAACGACGGCCAGT-3′) were custom-synthesized by Midland Certified Reagent using succinimidyl esters of the dyes. The dyes included tetramethylrhodamine (TMR), rhodamine green (RG), BODIPY-FL (BOD), and BODIPY-FL Br_2 (BBr), all from Molecular Probes, Eugene, OR, and Cy3 from Amersham Life Science, Inc., Cleveland, OH. BODIPY-R6G (BR6G)-labeled M13 primer was purchased from Molecular Probes. M13 forward primer labeled with the energy transfer BigDye, fluorescein−dTMR, was generously provided by Dr. L. G. Lee from PE Applied Biosystems. Solutions of the dye-labeled primers (0.4 μM) were prepared in 1× Tris-borate (TB) buffer (pH 8.4) with 7 M urea.

DNA Sequencing Sample Preparation. Cycle sequencing reactions were performed using standard cycle sequencing chemistry with AmpliTaq DNA polymerase (PE Applied Biosystems) and dye-labeled M13-21 primers on a pGEM-3Zf(+) double-strand DNA template (Promega, Madison, WI). A dye primer cycle sequencing core kit (PE Applied Biosystems) provided the reagents for the cycle sequencing reaction except for dye-labeled primers and template. Each termination reaction was performed using 4 μg pGEM-3Zf(+) double-stranded template DNA, 8 pmol of dye-labeled primer, 20 μL of termination mix, and 3 μL of AmpliTaq DNA polymerase in a total volume of 100 μL. The cycle protocol using a GeneAmp 2400 DNA thermal cycler (PE Applied Biosystems) was 15 cycles of 96 °C for 10 s, 55 °C for 5 s, and 70 °C for 60 s, followed by 15 cycles of 96 °C for 10 s and 70 °C for 60 s. Each termination sample was purified by ethanol precipitation. The sample was added into 800 μL of 95% ethanol, and the mixture was centrifuged at 20800*g* for 15 min. The supernatant was decanted, 1.5 mL of 70% ethanol was added to the precipitate, and the mixture was centrifuged for another 5 min. The supernatant was decanted and the DNA precipitate was allowed to dry. Prior to analysis, the sample was resuspended in 3−10 μL template suppression reagent and denatured by heating at 95 °C for 2 min.

Capillary Gel Electrophoresis. Separations were performed using linear polyacrylamide (LPA) gels. The internally coated, 100-μm internal diameter (i.d.) capillary was obtained from Beckman (Fullerton, CA). The replaceable LPA matrixes were prepared as follows: 0.4 or 0.5 g of acrylamide (ICN, Costa Mesa, CA) was dissolved in 10 mL of TB/7 M urea buffer (Beckman, Fullerton, CA) for preparing 4 or 5% LPA, respectively. The monomer solutions were polymerized by the addition at room temperature

(14) Metzker, M. L.; Lu, J.; Gibbs, R. A. *Science* **1996**, *271*, 1420.
(15) Seeger, S.; Bahteler, G.; Drexhage, K. H.; Arden-Jacob, J.; Deltau, G.; Galla, K.; Han, K. T.; Mueller, R.; Keollner, R. A.; Sauer, M.; Schulz, A.; Wolfrum, J. *Ber. Bunsen-Ges.* **1993**, *97*, 1542.
(16) Soper, S. A.; LeGendre, B. L. Jr.; Williams, D. C. *Anal. Chem.* **1995**, *67*, 4358.
(17) Bachteler, G.; Drexhage, K.-H.; Arden-Jacob, J.; Han, K.-T.; Kollner, M.; Muller, R.; Sauer, M.; Seeger, S.; Wolfrum, J. *J. Lumin.* **1994**, *60 & 61*, 511.
(18) Bachteler, G.; Drexhage, K.-H.; Arden-Jacob, J.; Han, K.-T.; Kollner, M.; Muller, R.; Sauer, M.; Seeger, S.; Wolfrum, J. *J. Lumin.* **1994**, *62*, 101.
(19) Flanagan, J. H., Jr.; Owens, C. V.; Romero, S. E.; Waddell, E.; Kajn, S. H.; Hammer, R. P.; Soper, S. A. *Anal. Chem.* **1998**, *70*, 2676
(20) Lieberwirth, U.; Arden-Jacob, J.; Drexhage, K. H.; Herten, D. P.; Muller, R.; Neumann, M.; Schulz, A.; Siebert, S.; Sagner, G.; Klingel, S.; Sauer, M.; Wolfrum, J. *Anal. Chem.* **1998**, *70*, 4771.
(21) Desilets, D. J.; Kissinger, P. T.; Lytle, F. E. *Anal. Chem.* **1987**, *59*, 1830.
(22) Li, L.; McGown, L. B. *Anal. Chem.* **1996**, *68*, 2737.
(23) Li, L.; He, H.; Nunnally, B. K.; McGown, L. B. *J. Chromatogr.* **1997**, *695*, 85.
(24) Nunnally, B. K.; He, H.; Li, L.; Tucker, S. A.; McGown, L. B. *Anal. Chem.* **1997**, *69*, 2392.
(25) He, H.; Nunnally, B. K.; Li, L.; McGown, L. B. *Anal. Chem.* **1998**, *70*, 3413.

of 2 μL of 10% (w/v) ammonium persulfate (ICN) and 1 mL of 100% TEMED (ICN) per mL of monomer solution to yield final concentrations of 0.02% (w/v) APS and 0.1% (v/v) TEMED.[26]

The LPA matrix was replaced in the capillary before each run by using a 50-μL modified HPLC syringe (Hamilton Co., Reno, NV). The capillary then was electrodialyzed for 40 min at an electric field of 200 V/cm in order to decrease the background fluorescence. CE separation of the products of the DNA cycle sequencing reactions was performed using an electric field strength of 200 V/cm and a capillary that was typically 57 cm long and 44 cm to the detection window.

On-the-Fly Fluorescence Lifetime Detection. The MHF fluorescence lifetime instrument (Model 4850MHF or 48000MHF, Spectronics Instruments, Rochester, NY) was interfaced to a Beckman P/ACE 5000 CE system for on-the-fly detection as previously described.[15,23,25] The CE system was equipped with a CE/MS interface to provide an external power supply to accommodate the interface to the MHF. The excitation source was either an air-cooled argon ion laser (model 543R-AP-A01, Omnichrome, Chino, CA), which provided 100−125 mW at 488 or 514 nm, or a water-cooled argon ion laser (Coherent model 307), which provided 100−300 mW at 488 or 514 nm. The laser beam was focused onto the detection window of the capillary using either a 45-mm focusing lens or a 6.3× microscope objective having a focal length of 22 mm. The emission signal was collected by a 40× (Oriel, Stratford, CT) or a 63× (Spindler and Hoyer)microscope objective. For 488 nm excitation, incident laser light was passed through a 488-nm band-pass filter (Oriel) and emission was selected through a 515-nm long pass filter (Oriel) combined with a 488-nm holographic filter (Notch-Plus, Kaiser Optical Systems, Ann Arbor, MI) to further reduce the contribution from scattered laser light. For 514-nm excitation, emission was selected through a 550-nm long pass filter (Oriel).

A cross-correlation frequency of 10 or 9.4 Hz was used in the on-the-fly lifetime measurements, which resulted in 10 or 9.4 phase and modulation measurements per second. Ten successive measurements were then averaged prior to data analysis to yield approximately one lifetime measurement per second. Scattered light from the capillary provided the lifetime reference.

Because of the nature of the on-the-fly lifetime detection technique used in this work, it was not possible to extract meaningful standard deviations associated with each lifetime measurement. Instead, standard deviations were calculated from separations of dye-labeled fragments for each individual base by averaging the lifetimes across the CE peaks. The lifetimes are reported as these average values. The standard deviations typically were in the range of 1−2%.

Data Analysis. In on-the-fly detection, both fluorescence intensity and lifetime are recovered from the same dynamic MHF data to provide intensity-lifetime electropherograms. The fluorescence intensity was recovered from the dynamic MHF data using an in-house program.[27]

Two methods were used for base calling. In the first, bases were identified directly from the lifetimes recovered from one-component nonlinear least-squares (NLLS) analysis of the lifetime data. The recovered lifetimes were compared to predetermined lifetimes obtained for sequencing products of each individual base under identical conditions. In regions of poor electrophoretic resolution, the application of this method may result in inaccuracies due to multiexponential decay from overlapping peaks.

In the second method, lifetime-resolved electropherograms were obtained for each terminal base using NLLS analysis of the on-the-fly lifetime data in which the lifetime components in the fitting models were fixed to predetermined values that were obtained for sequencing products of each individual base under identical conditions. The recovered fractional intensity of each lifetime component at each point in the electropherogram of the mixture was then multiplied by the total intensity at that point to reconstruct the electropherogram for each of the individual bases in the mixture.

RESULTS AND DISCUSSION

Optimization of Experimental Conditions. The detection of products of DNA sequencing reactions is a challenge for on-the-fly fluorescence lifetime detection due to the low concentrations of the DNA fragments. In DNA sequencing reactions, 0.4 μM dye-labeled primer was used in each reaction to generate labeled sequencing fragments. Because the dye-labeled primer is in excess in order to maximize the yield of dye-labeled fragments, only a fraction of the dye-labeled primer is actually extended and used to label sequencing fragments. The concentration of the dye is further diluted by the hundreds of fragments that are generated in each sequencing reaction, which results in very low concentrations of each labeled fragment.

Some improvements of our on-the-fly lifetime detection were made to meet the challenge of low concentrations of labeled fragments. First, it was desirable to use a microscope objective having higher magnification power and a large numerical aperture (N. A.) to collect more of the emitted light.[28−30] The 10× objective with 0.25 N. A. used in previous studies of dye-labeled primers[15,24,25] was replaced first by a 40× objective with 0.65 N. A., which should increase the collection efficiency by a factor of 6.3.[28] The efficiency can be increased by an additional factor of 2 by using a 60× objective, but most 60× objectives (0.85 N. A.) have working distances that are too short for our 375-μm o.d. capillary. Fortunately, we were able to use a 63× objective that had enough working distance to accommodate the thickness of our capillary.[28]

Because of the small inner diameter of the capillary, the incident laser beam must be passed through a focusing device. A 45-mm focusing lens was used for the initial experiments in this work. It was later replaced with a 6.3× microscope objective with a 22-mm focal length in order to focus the laser beam more tightly onto the inner separation channel of the capillary.

The CE injection method can affect the separation resolution. For DNA sequencing, electrokinetic injection must be used because pressure injection will disrupt the gel matrix. The amount of sample injected is determined by the injection voltage and the injection time. Although lower injection voltages with longer injection times can inject the same amount of sample, some

(26) Ruiz-Martinez, M. C.; Berka, J.; Belenkii, A.; Foret, F.; Miller, A. W.; Karger, B. L. *Anal. Chem.* **1993**, *65*, 2851.
(27) Smalley, M. B.; Shaver, J. M.; McGown, L. B. *Anal. Chem.* **1993**, *65*, 3466.
(28) Wu, S.; Dovichi, N. J. *J Chromatogr.* **1989**, *480*, 141.
(29) Hernandez, D.; Marquina, R.; Escalona, J.; Guzman, N. A. *J. Chromatogr.* **1990**, *502*, 247.
(30) Hernandez, D.; Escalona, J.; Joshi, N.; Guzman, N. *J. Chromatogr.* **1991**, *559*, 183.

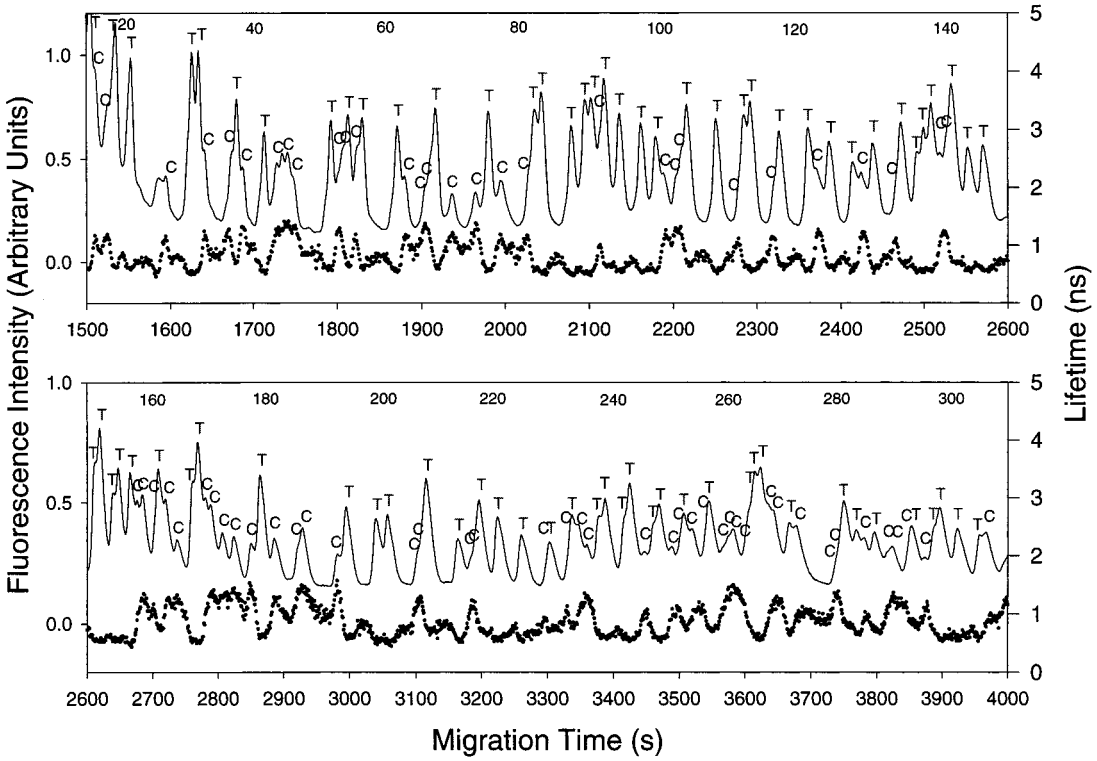

Figure 1. Fluorescence lifetime (dotted line) and intensity (solid line) electropherograms of a mixture of TMR-labeled, C-terminated fragments and Cy3-labeled, T-terminated fragments using 514-nm excitation and 5% LPA separation. Lifetimes were recovered from NLLS fits to a one-component model. Injection: 10 kV × 20 s.

researchers suggested that it might give a different separation resolution.[31−33] In a separation of RG-labeled, G-terminated fragments, we obtained better resolution and more symmetrical peaks using a 2.5 kV × 80-s injection, as compared to a 10 kV × 20-s injection.

514 nm Dye System. The 514-nm dye system initially included Cy3, BBr, TMR, and BODIPY530/550 (B530). The results from on-the-fly fluorescence lifetime detection of labeled primers showed that these four dyes have well-separated lifetimes, but BBr and B530 are weaker than the other dyes, which limits detectability of the labeled sequencing products. B530 was replaced with BR6G, which has similar spectral and lifetime properties but is more strongly fluorescent. On-the-fly lifetimes of the dye-labeled sequencing products were 0.5 ns for Cy3-labeled, T-terminated fragments; 2.1 ns for TMR-labeled, C-terminated fragments; and 5.0 ns for BR6G-labeled, T-terminated fragments. A fourth dye to replace BBr could not be found, however, and we eventually focused on the 488-nm system described below. Nevertheless, studies of a two-component system are described here to illustrate the performance of multiplex, on-the-fly detection of sequencing products at 514 nm in hopes that a fourth dye may eventually be found.

A mixture of TMR-labeled, C-terminated fragments and Cy3-labeled, T-terminated fragments was separated using 5% LPA gel and a 10 kV injection for 20 s and was detected by on-the-fly fluorescence lifetime detection using a 514-nm laser line (125 mW) for excitation. The results recovered from NLLS fits to a one-component model are shown in Figure 1. The recovered lifetimes of the Cy3-labeled, T-terminated fragments and the TMR-labeled, C-terminated fragments were 0.5 ns and 1.5 ns, respectively. The lifetimes of the TMR-labeled fragments recovered from the mixture were lower than the values given above for the individual TMR-labeled fragments. This is most likely due to the influence in the one-component fits of the more intense, shorter-lifetime Cy3 peaks that overlap with the TMR peaks across the entire electropherogram.

The assignments of T and C directly from the recovered lifetimes are marked in Figure 1. From 16 to 320 bases, 9 errors of base calling were found, including one T reading error due to abnormal enzymatic behavior (an extra fragment was generated during the sequencing reaction). The accuracy of the base calling was 94%. Most reading errors were associated with severely overlapping peaks.

Base calling was also performed using the lifetime-resolved intensity electropherograms of TMR-labeled, C-terminated fragments and Cy3-labeled, T-terminated fragments. This data analysis scheme has the advantage of being able to resolve overlapping peaks so that base calling can be done even for severely overlapping peaks. The electropherograms were resolved using NLLS analysis in which the lifetimes of TMR-labeled fragments and Cy3-labeled fragments were fixed to the predetermined values recovered from injections of the individual sequencing reaction products.

Using the recovered electropherograms (not shown), we found only one base-calling error in the region of 16−320 bases, giving

(31) Paulus, A.; Ohms, J. L. *J. Chromatogr.* **1990**, *507*, 113.
(32) Demorest, D.; Dubrow, R. *J. Chromatogr.* **1991**, *559*, 43.
(33) Salas-Solano, O.; Ruiz-Martinez, M. C.; Carrilho, E.; Kotler, L.; Karger, B. L. *Anal. Chem.* **1998**, *70*, 1528.

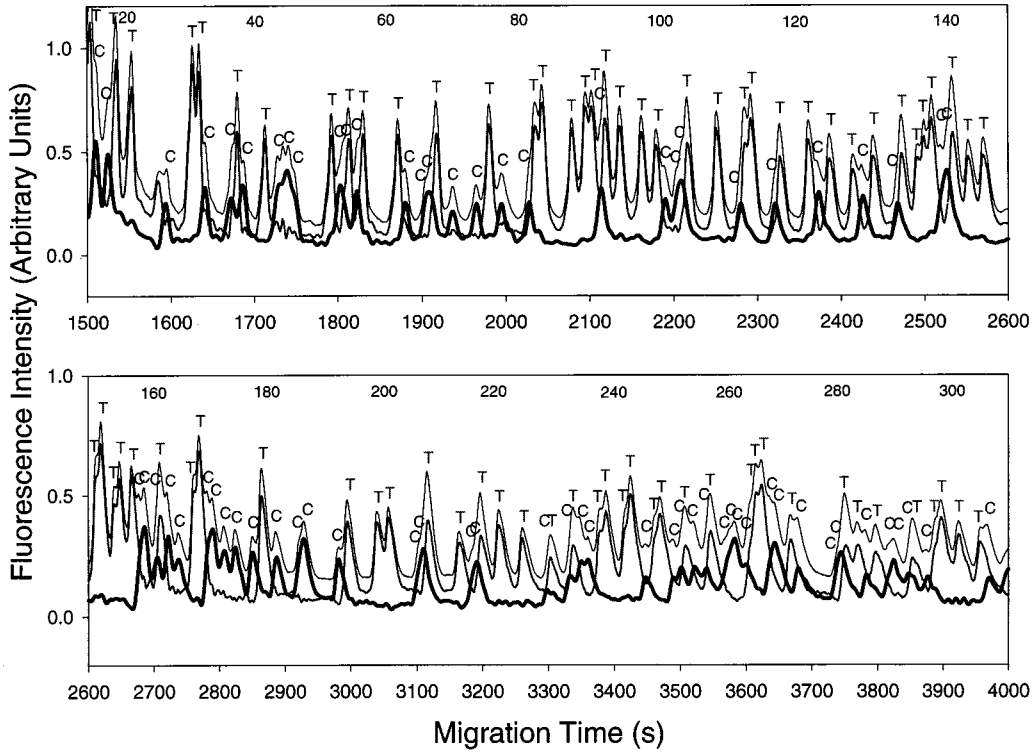

Figure 2. Lifetime-resolved intensity peaks of individual dye-labeled DNA sequencing fragments recovered from the same run used for Figure 1. Peaks were recovered from two-component NLLS analysis in which the lifetimes were fixed to predetermined values. The recovered intensity electropherograms of C- and T-terminated fragments were smoothed by removing high-frequency noise.

a base-calling accuracy of >99%. The single error was the 26-base fragment migrating at 1590 s that was attributed to abnormal enzymatic behavior. Above 320 bases, base calling became difficult due to decreased separation efficiency. Three errors were found in the region from 320 to 360 bases.

To improve the accuracy of base-calling from the lifetime-resolved electropherograms, data smoothing was performed to remove high-frequency noise from the lifetime analysis. The recovered electropherograms were first transformed to frequency-domain data using fast Fourier transform in MATLAB. High-frequency noise was then removed and a reverse fast Fourier transform was performed to transform the truncated frequency-domain data back to time-domain data. It should be noted that similar results can be obtained by using a low-pass frequency filter in MATLAB.

Figure 2 shows the lifetime-resolved electropherograms of TMR- and Cy3-labeled fragments after the removal of high-frequency noise. The assignments of C and T are marked on the figure. The smoothed data was easier to read because the noise introduced from the lifetime data analysis was removed, and the features of the peaks were more obvious. From 16 to 320 bases, the base-calling accuracy was the same as before; the only error was due to the ghost peak from abnormal enzymatic behavior. From 320 to 360 bases, the reading was more accurate, with only one error. The overall base-calling accuracy from 16 to 360 bases was 99%. We conclude that base-calling accuracy is significantly improved by using lifetime-resolved electropherograms instead of direct lifetimes obtained from one-component NLLS analysis.

488-nm Dye System. The original 488-nm dye system that was proposed for four-decay detection included NBD-aminohexanoic acid (NBD), TMR, RG, and BODIPY-FL (BOD).[24] These dyes were selected on the basis of the lifetimes of the free dyes in batch solution. It was subsequently discovered that the lifetime of TMR-labeled primer is different from that of free TMR and indistinguishable from the lifetime of RG-labeled primer. TMR was, therefore, replaced by Cy3, which has suitable intensity and lifetime. It was then necessary to replace NBD because it is a relatively weak dye, and its lifetime is similar to that of Cy3.

A three-component mixture of Cy3-labeled, T-terminated fragments, RG-labeled, C-terminated fragments, and BOD-labeled, G-terminated fragments was separated using 4% LPA gel and detected using 488-nm excitation. The lifetime results recovered using one-component NLLS analysis are shown in Figure 3. Figure 4 shows an expanded view from 2800 s to 3300 s. The base calling of T, C, and G was based on the lifetimes of the peaks recovered from the one-component fit. The Cy3-labeled, T-terminated fragments have lifetimes of 0.9 ns; the RG-labeled, C-terminated fragments have lifetimes of 2.3 ns; and BOD-labeled, G-terminated fragments have lifetimes of 3.3 ns. From 51 to 300 bases, 19 base-calling errors were found, which is equivalent to an accuracy of 90%. The gel gave poor separation for DNA fragments shorter than 50 bases and longer than 300 bases.

Base calling was also performed using the lifetime-resolved intensity electropherograms for each of the bases. Peaks were resolved using three-component NLLS analysis in which the lifetimes were fixed to the predetermined values. The use of resolved peaks extended the base-calling range to 320 bases. Using the raw recovered electropherograms of various dye-labeled fragments (not shown), 6 reading errors were found, which is equivalent to an accuracy of 97%. Base calling from the smoothed electropherograms (shown in Figure 5) gave only 3 errors, which is equivalent to an accuracy of 98.5%. The improved read length

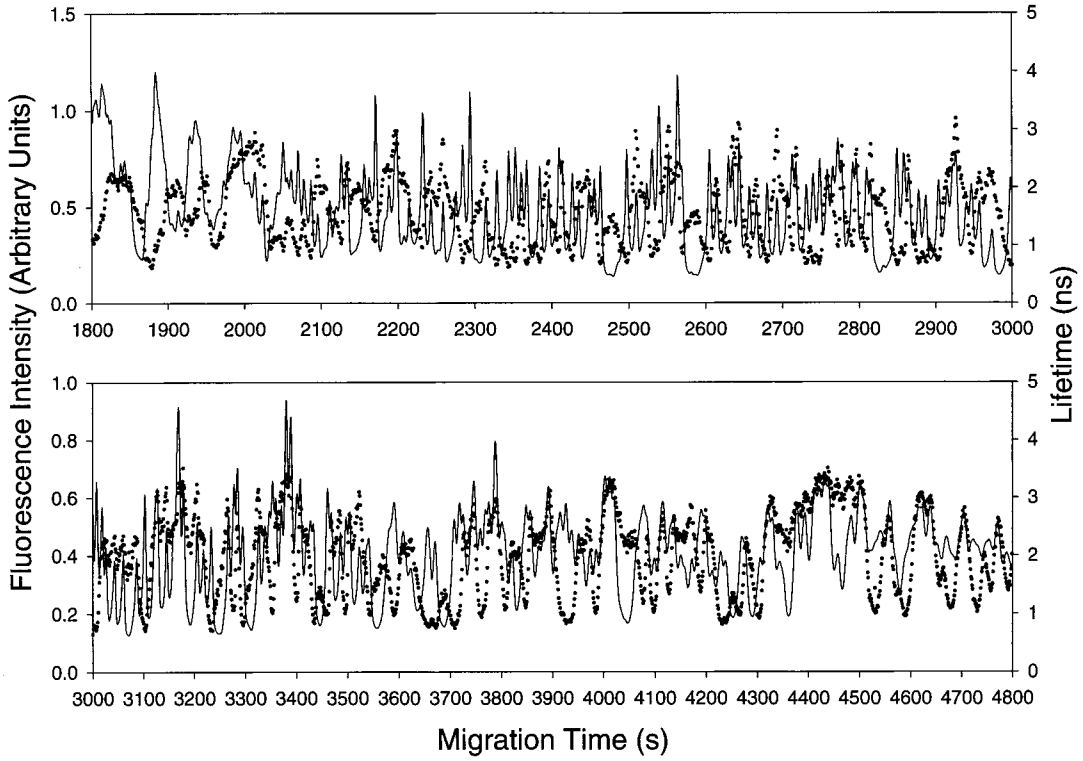

Figure 3. Fluorescence lifetime (dotted line) and intensity (solid line) electropherograms of a mixture of Cy3-labeled, T-terminated fragments; RG-labeled, C-terminated fragments; and BOD-labeled, G-terminated fragments using 488-nm excitation and 4% LPA separation. Lifetimes were recovered from NLLS fits to a one-component model. Injection: 10 kV × 20 s.

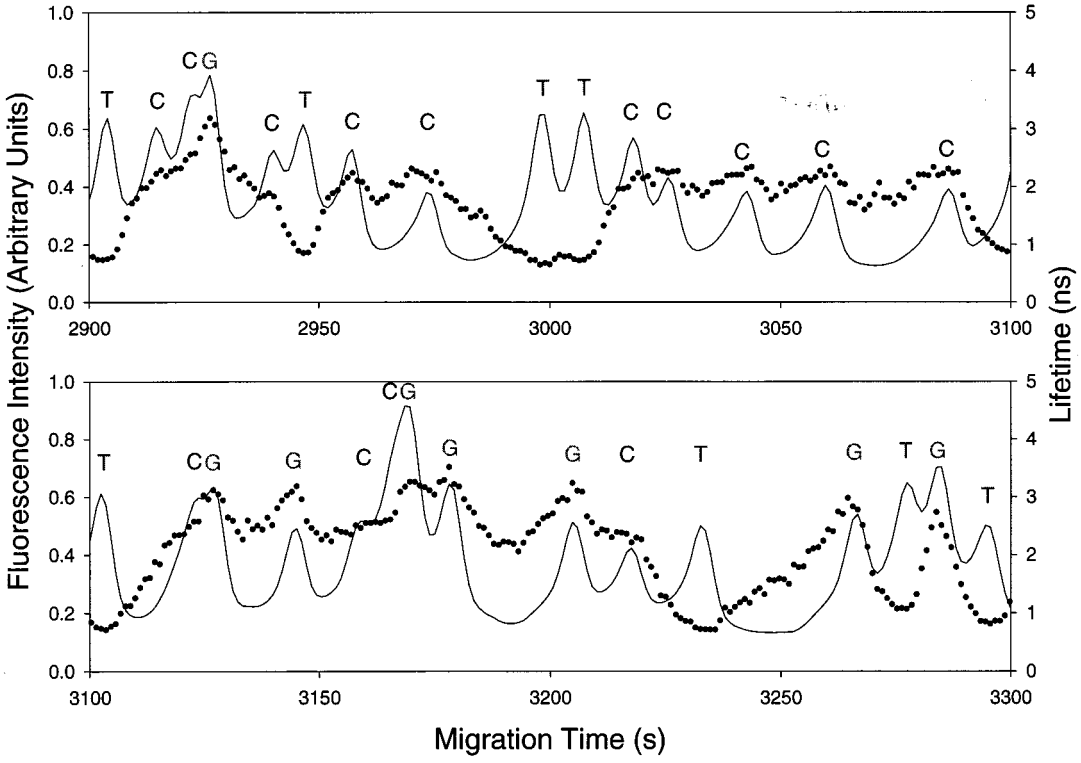

Figure 4. Expanded scale of a section of the fluorescence lifetime (dotted line) and intensity (solid line) electropherograms in Figure 3.

and accuracy are due to improved base-calling for severely overlapping peaks.

DNA Sequencing Using Sequential Injections of Dye-Labeled Sequencing Reaction Products. The most recent development in the 488-ns dye system is the replacement of NBD by fluorescein−dTMR, an energy transfer BigDye[12,13] that contains the 5-carboxy isomer of 4′-aminomethylfluorescein as the donor dye and dTMR as the acceptor dye. The dTMR is 4,7-dichloro-

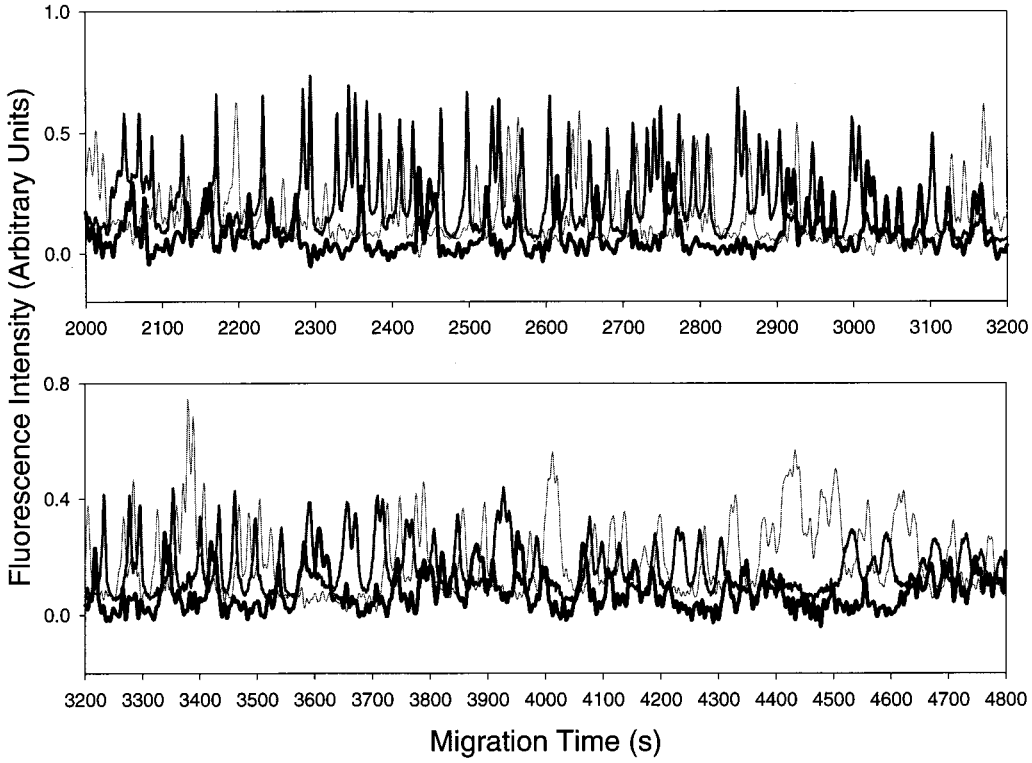

Figure 5. Lifetime-resolved intensity peaks of individual dye-labeled DNA sequencing fragments from a mixture of Cy3-labeled, T-terminated fragments (dark gray); RG-labeled, C-terminated fragments (black); and BOD-labeled, G-terminated fragments (light gray). Peaks were recovered from three-component NLLS analysis in which the lifetimes were fixed to predetermined values. The recovered intensity electropherograms of T-, C-, and G-terminated fragments were smoothed by removing high-frequency noise.

substituted TMR that has a narrower emission spectrum than TMR. A rigid amino acid linker, 4-aminomethylbenzoic acid, separates the dyes in the BigDye molecule. Fluorescein−dTMR has high energy transfer efficiency when excited at 488 nm. There was essentially no interference from the donor in four-decay detection.

The 488-nm dye system of Cy3, fluorescein−dTMR, RG, and BOD provided good lifetime resolution, but successful implementation requires mobility correction for the fluorescein−dTMR-labeled fragments because the dye is much larger than the three other dyes. One approach is to use a longer linker between dye and DNA for the other three dyes in order to decrease the migration speed of their dye-labeled fragments.[14,21]

We used a different approach to compensate for the mobility difference. Sequential injections of the four sets of dye-labeled sequencing reaction products were performed in which the fluorescein−dTMR-labeled fragments were injected first and allowed to migrate for a short time before sequentially injecting the other dye-labeled fragments.

There are several advantages of the sequential approach when compared to simultaneous injection of the four sets of fragments as a mixture. Separation resolution was improved because the injection time for each dye-labeled fragment was shortened to about one-fourth of that for the mixture, and intensity matching among the four dye-labeled fragments could be done by controlling the injection time for each set of fragments. However, becasue the proportional effect of the dye on mobility decreases as fragment length increases, sequential injection can only compensate for the mobility difference within a certain range of the separation.

Figure 6 shows the sequencing results using sequential injections of A (fluorescein−dTMR), C (RG), T (Cy3), and G (BOD) fragments onto a 5% LPA-packed column. The injection sequence was a 2.5 kV × 20-s injection of fluorescein−dTMR-labeled A-terminated fragments; followed after 70 s by a 2.5 kV × 10-s injection of RG-labeled, C-terminated fragments; a 2.5 kV × 20-s injection of Cy3-labeled, T-terminated fragments; and finally, a 2.5 kV × 40-s injection of BOD-labeled, G-terminated fragments. Figure 7 shows an expanded view of the results from 3500 to 3900 s. Mobility correction was successful within this range of fragment length. Cy3-labeled, T-terminated fragments have lifetimes of 1.7 ns; fluorescein−dTMR-labeled, A-terminated fragments have lifetimes of 2.5 ns; RG-labeled, C-terminated fragments have lifetimes of 2.9 ns; and BOD-labeled, G-terminated fragments have lifetimes of 3.5 ns.

Base calling from 41 to 220 bases was based directly on lifetimes of the peaks. Eight errors occurred, which is equivalent to an accuracy of 96%. Base calling using lifetime-resolved peaks was unsuccessful due to difficulties in four-component NLLS analysis, even when the lifetimes were fixed to predetermined values. Improved approaches are under investigation for data analysis of four-component systems which, in our experience, is generally problematic for any real system with noise and background.

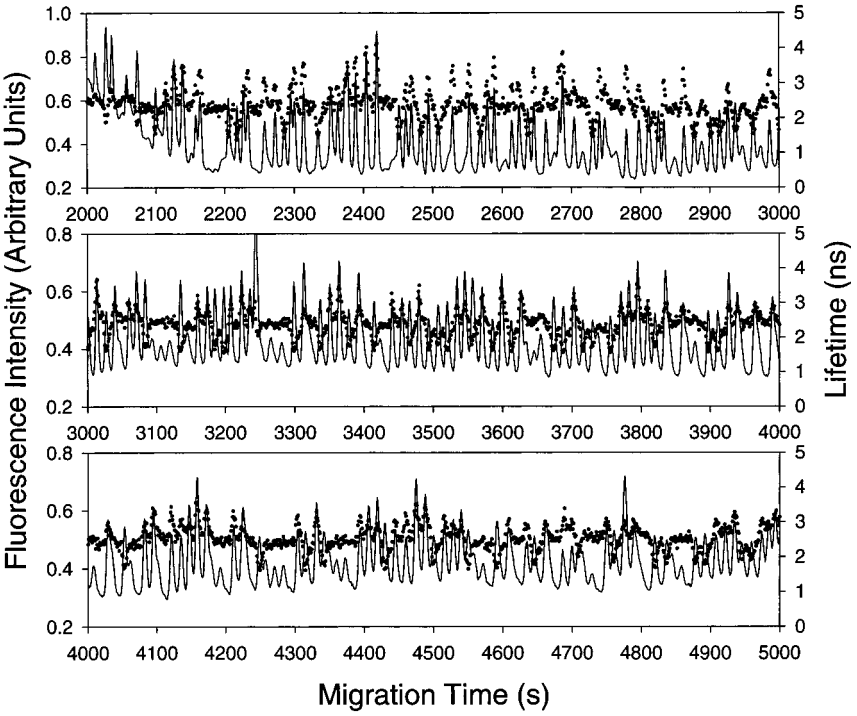

Figure 6. Fluorescence lifetime (dotted line) and intensity (solid line) electropherograms of sequential injections of A (fluorescein−dTMR), C (RG), T (Cy3), and G (BOD) fragments. Injection sequence: A (2.5 kV × 20 s), 70-s separation, C (2.5 kV × 10 s), T (2.5 kV × 20 s), G (2.5 kV × 40 s) using 488-nm excitation and 5% LPA separation. Lifetimes were recovered from NLLS fits to a one-component model.

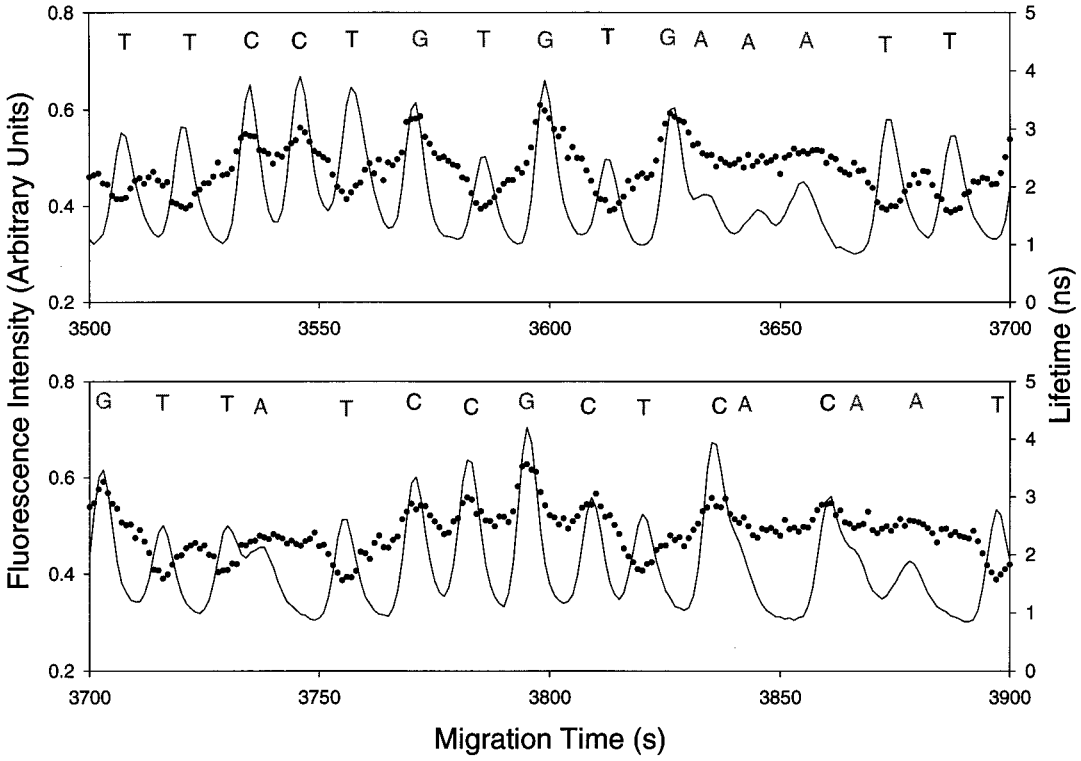

Figure 7. Expanded scale of one section of the fluorescence lifetime (dotted line) and intensity (solid line) electropherograms in Figure 6.

CONCLUSIONS

This work demonstrates frequency-domain, on-the-fly fluorescence lifetime detection and identification of dye-labeled DNA sequencing reaction products for DNA sequencing. Two four-dye systems, one excited at 488 nm and the other at 514 nm, were investigated. The 514-nm dye system provided good lifetime resolution but was difficult to apply to four-decay detection because one of the dyes, BBr, is a very weak dye. The 488-nm dye system, including (in the order of increasing lifetime) Cy3, fluorescein−dTMR, RG, and BOD, provided good lifetime resolution and strong

signal intensity, but required compensation for mobility differences due to the large size of fluorescein−dTMR. This was accomplished by using sequential injections of the sequencing reaction products of each of the individual bases.

Two base-calling methods were applied: direct identification of the peaks based on fits to a one-component lifetime model and identification of peaks from lifetime-resolved electropherograms of each of the four sets of dye-labeled fragments. The first method is simpler to apply but less accurate; the second method is more accurate but difficult to apply to the four-component system due to limitations in analysis of the lifetime data. The first method is effective if there is moderate overlap, but the second method is preferable because it extends the read length to longer fragments where the separation efficiency of the column is reduced.

Base-calling accuracy in a three-component separation was improved from 90% for the first method to 98.5% for the second method that was obtained for a longer read length. Accuracy was further improved by removing noise from reconstructed electropherograms for each of the different bases. Base-calling accuracies of 99% from 16 to 360 bases, 98.5% from 51 to 320 bases, and 96% from 41 to 220 bases were obtained for mixtures of sequencing products for two bases, three bases, and all four bases, respectively.

Clearly, the use of lifetime-resolved intensity electropherograms significantly improves the base-calling accuracy. This approach has not yet been successful for four-component mixtures of sequencing products due to data analysis limitations. Noise and uncertainty in the four-component data overwhelm the ability of NLLS analysis to extract four components, even when the lifetimes of the components are well separated from each other and predetermined. New strategies are needed to overcome these limitations. Improvements in the quality of data through instrumental modifications could also be helpful; however, the excellent results that were obtained for the three-component system suggest that data analysis is the main weakness because the quality of the data for the three- and four-component systems was similar.

ACKNOWLEDGMENT

This work was supported by the National Institutes of Health (Grant R01HG01161). We are grateful to Dr. L. G. Lee from PE Applied Biosystems for generously providing M13 forward primer labeled with fluorescein−dTMR.

Received for review August 11, 2000. Accepted October 14, 2000.

AC000952N