Nucleic Acids Research, 1993, Vol. 21, No. 4 803–806

Rapid sizing of individual fluorescently stained DNA fragments by flow cytometry

Peter M.Goodwin, Mitchell E.Johnson, John C.Martin[1], W.Patrick Ambrose, Babetta L.Marrone[1], James H.Jett[1] and Richard A.Keller*

Chemical and Laser Sciences Division and [1]Life Sciences Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA

Received January 15, 1993; Accepted January 22, 1993

ABSTRACT

Large, fluorescently stained restriction fragments of lambda phage DNA are sized by passing individual fragments through a focused continuous wave laser beam in an ultrasensitive flow cytometer at a rate of 60 fragments per second. The size of the fluorescence burst emitted by each stained DNA fragment, as it passes through the laser beam, is measured in one millisecond. One hundred sixty four seconds of fluorescence burst data allow linear sizing of DNA with an accuracy of better than two percent over a range of 10 to 50 kbp. This corresponds to analyzing less than 1 pg of DNA. Sizing of DNA fragments by this approach is much faster, requires much less DNA, and can potentially analyze large fragments with better resolution and accuracy than with gel-based electrophoresis.

INTRODUCTION

Gel electrophoresis is the most widely accepted technique for analysis and separation of DNA fragments. Over the years, a wide variety of approaches to gel electrophoresis have been developed that have improved or enhanced the resolution or separation capabilities (1). Standard gel electrophoresis is used for fragment sizes up to approximately 50 kb in length. Larger fragments must be separated by some form of pulsed field electrophoresis (2). Capillary gel electrophoresis (3) and ultrathin slab gel electrophoresis (4) are currently being developed to allow for high speed separation of DNA sequencing ladders for sizes less than 1 kb. No matter what form of electrophoresis is used, the separation is highly non-linear and generally has an upper limit to the size of the DNA that can be analyzed. In addition, most gel-based separations take many hours, depending on required resolution and detection method.

In this paper we report on a new approach to sizing, and ultimately separating, DNA fragments. The technique is based upon flow cytometry and results in the measurement of DNA fragment sizes that is linear in both the size calibration and in the response to the number of fragments measured. The linearity of the size measurement follows from the fact that what is measured is the fluorescence from stoichiometrically stained individual fragments of DNA as they flow through a laser beam. The linearity of counting the number of fragments of each size follows since each fragment is measured individually.

We demonstrate that large, individual, fluorescently stained DNA fragments can be sized accurately at a rate of approximately 60 fragments per second (5). A histogram of the fluorescence intensities of individual fragments (10 to 50 kbp) in a mixture of lambda phage DNA (λ DNA), a *Kpn* I digest of λ DNA, and an *Apa* I digest of λ DNA provides linear sizing with an accuracy of better than 2% in 164 seconds. Detection and analysis of individual fluorescently stained DNA fragments in the one millisecond that they are in the laser beam requires sensitive fluorescence detection capabilities. Our apparatus is a modification of the instrumentation we used to detect single dye molecules and to measure their fluorescence lifetimes (6, 7).

MATERIALS AND METHODS

DNA staining

A mixture of λ DNA (48,502 bp), a *Kpn* I digest of λ DNA (29,946 bp, 17,053 bp, and 1,503 bp), and an *Apa* I digest of λ DNA (38,416 bp and 10,086 bp) was stained stoichiometrically with the DNA intercalating dye TOTO-1 (8) in the following manner. Lambda DNA and the *Kpn* I and *Apa* I digests of λ DNA were stained separately with TOTO-1 dye at a DNA concentration of 400 pg/μl for 60 minutes. The TOTO-1 concentration was 1.2×10^{-7} M, to yield an average bp:dye ratio of 5. The analyzed mixture was prepared by adding equivalent aliquots of each stained DNA solution to buffer diluent for a final total fragment concentration of approximately 10^{-13} M, or about 3 ng/ml. All solutions were prepared in pH 8 TE buffer (10 mM Tris–HCl, 1 mM EDTA). Staining solutions were prepared fresh daily and all TOTO-1 dilutions were prepared immediately before staining. TOTO-1 , a thiazol orange homodimer, was obtained from Molecular Probes, Eugene, OR; λ DNA was obtained from GIBCO BRL, Gaithersburg, MD; and the λ DNA digests were obtained from New England Biolabs, Beverly, MA.

* To whom correspondence should be addressed

Ultrasensitive flow cytometry apparatus

Flowing, fluorescently stained fragments were hydrodynamically focused in a modified flow chamber. The flow cell was a modified commercial square bore sheath flow cuvette (Ortho Diagnostics, Inc.; currently available from BDIS) with a 250 μm internal dimension. The collection side was ground down to a cover slip thickness to allow positioning of the collection microscope objective at its working distance from the sample stream (see below). The sheath fluid was water and the flow was produced by gravity feed at approximately 30 μl/min. The sample flow was also produced by gravity feed at rates less than 1 μl/min, or a linear velocity of approximately 5 cm per second. The sample flowed at a rate of approximately 60 fragments per second. The average transit time (time spent in the laser beam) of each fragment was 870 μs, as determined by autocorrelation data (9). The focused sample stream diameter was approximately 20 μm. Under these conditions, the probability of two or more fragments being in the probe volume at the same time is less than 0.5%.

Thirty milliwatts of continuous wave (cw) Ar^+ laser radiation (Coherent 171) at 514.5 nm was focused with a spherical plano convex lens to a beam waist of 40 μm ($1/e^2$ diameter) in the center of the flow chamber. Fluorescence from the intercalated dye was collected by a 40×, 0.85 NA microscope objective (Nikon Fluor). A 1 mm slit oriented parallel to the flow axis and located at the image plane of the objective served as a spatial filter. The probe volume thus defined was approximately 130 pl. The light was then spectrally filtered to remove Rayleigh and Raman scatter from the water solvent. The filter pack consisted of a Raman holographic edge filter (RHE514.5, Physical Optics Corporation) and a bandpass filter centered at 530 nm with a 30 nm passband (530DF30, Omega Optical) for a detection bandwidth of 520−545 nm. The transmitted fluorescence was focused on the photocathode of a cooled photomultiplier tube (Hamamatsu R636) operating in the photon counting mode. The overall detection efficiency (photoelectrons counted divided by photons emitted) was approximately 2×10^{-3}.

Photoelectron pulses (PE) from the photomultiplier tube were amplified (Phillips Scientific 6954-S50), conditioned with a constant fraction discriminator (Tennelec TC454), and counted on a multichannel scaler (MCS, Lecroy 3521A/MM8201 A). The channel address of the MCS was advanced by a clock that ran at 10 kHz (100 μs wide bins). The CAMAC counting electronics were interfaced to a computer workstation (SUN, SPARC station IPC) via GPIB. Data analysis was carried out on the SPARC station.

The raw MCS data stream was sifted for bursts in the following manner. The average background rate was determined from the data during quiescent periods between fluorescence bursts. After subtraction of the average background, contiguous (in time) data points above a predetermined threshold were accumulated to give the burst size in photoelectrons. The average background rate for the data shown below was 20 photoelectrons per bin; the threshold was set at 25 photoelectrons per bin.

In order to monitor the performance of the cytometer, yellow-green fluorescent microspheres were used as standards. The spheres had a specified mean diameter of 0.913 μm with a standard deviation of the diameter of 0.03 μm (10). The microspheres were obtained from Polyscience, Inc. In our apparatus, a neutral density filter of OD 3 was used to attenuate the fluorescence from the spheres to the level of signal from stained λ DNA . The CV of the fluorescence of the spheres was measured independently on a different flow cytometer (11), and was found to be 0.022 at an excitation wavelength of 514.5 nm.

Gel electrophoresis

Lambda DNA and the digests of λ DNA were also analyzed by gel electrophoresis. The analysis was run on a 30 cm, 0.5% agarose gel at 40 V, 0.09 A for 22 hr. Lanes were loaded with 100 ng of λ DNA, 200 ng of digest, and a tracking dye. Staining was accomplished by adding ethidium bromide to the electrophoresis buffer. See Figure 3 caption for details of lane composition.

RESULTS AND DISCUSSION

The histogram shown in Figure 1 is the result of analyzing 164 seconds of data containing measurements of approximately 10,000 DNA fragments (approximately 0.3 pg of DNA). The 48, 38, 30, 17, and 10 kbp fragments are well resolved. The signature of the 1.5 kbp fragment is masked by background due to residual scattered light and impurity fluorescence. Centroids of the peaks in the histogram were obtained by fitting the data to the sum of five Gaussians plus an exponential background. The reduced χ^2 of the fit is 0.999. The results of the fit are summarized in Table 1 and the fitted curve is shown in Figure 1. The coefficient of variation (CV) (10) of the individual, separated peaks in the histogram ranges from 5−14%. A CV of 5% was also obtained from the 1 μm fluorescent microspheres, the fluorescence intensity of which was adjusted to be the same as that of λ DNA. The larger CVs obtained for the smaller fragments, as shown in Table 1, result from decreasing signal to noise ratio of the burst data and proportionately higher shot noise. Our CVs are within a factor of 2 of the shot noise limit. Our current resolution at 48.5 kbp approaches that of gel electrophoresis. For example, analysis of data from Reese and Zimm yields a CV of 4% for gel electrophoresis of T7 DNA (40 kbp) (12). We expect to reduce the CV of our measurements to about 1% (see below).

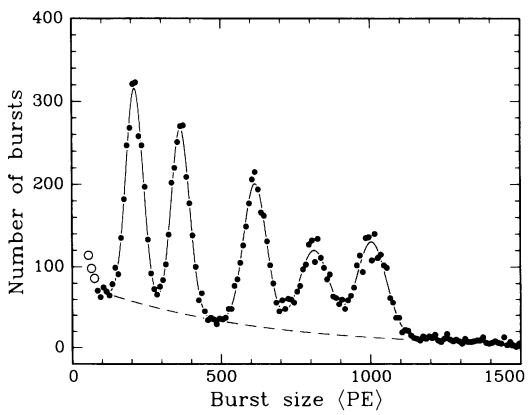

Figure 1. Histogram of the fluorescence burst sizes (photoelectrons counted) from a mixture of λ DNA, a *Kpn* I digest of λ DNA, and an *Apa* I digest of λ DNA containing fragment sizes of 1,503 bp; 10,086 bp; 17,053 bp; 29,946 bp; 38,416 bp; and 48,502 bp. The fragment at 1,503 bp is masked by scatter and impurity fluorescence backgrounds. The circles represent measured data and the solid line is the fit to the data. The three open circles were not included in the fit. The dashed line is the exponential fit to the underlying background. The histogram bin width is 5 PE.

A plot of the centroids of the peaks obtained from the fit verses the known DNA fragment lengths is shown in Figure 2. A linear least squares fit to the data is also shown on the plot. The correlation coefficient of the fit is R=0.9996. The fitted line passes through the origin within the accuracy of the fit (intercept is +5 ± 10 PE). Deviations of the measured values from the fitted line are listed in Table 1 and are less than 2% with an average deviation of 1.4% (13). Results were reproducible from run to run, but day to day reproducibility depended critically on instrument alignment. and the use of freshly stained DNA (within 12 hours). Preliminary results were obtained for ethidium homodimer stained DNA. In this case, however, the reduced signal and the increased water Raman background present in the detection bandwidth required the use of picosecond pulsed excitation and time gated detection (14).

To verify the identity of the peaks shown in Figure 1, λ DNA and the λ DNA digests were also sized using conventional gel electrophoresis. Figure 3 shows a photograph of the gel. The peaks identified in Figure 1 clearly correspond to the bands in lane 'f' of the gel. The remaining lanes 'a−e' (i.e. individual digests and λ DNA) are shown to reaffirm the identity of the bands in lane 'f'. Note that the 1.5 kbp fragment, which was masked by background in the measurement, though faintly visible on the gel, did not reproduce well in the photograph. The gel electrophoresis results demonstrate that the analysis mixture used to produce the histogram contained only fragments of the sizes claimed by the manufacturer.

Table 1. Data analysis of the histogram in figure 1

Fragment Size (bp)	Burst Centroid (PE)	δ	CV	√n/n
10086	210.8	0.016	0.137	0.068
17053	364.6	−0.017	0.084	0.053
29946	614.8	0.017	0.063	0.041
38416	813.0	−0.016	0.064	0.035
48502	1004.1	0.005	0.050	0.032
Fluorescent microspheres	955.9		0.046	0.032

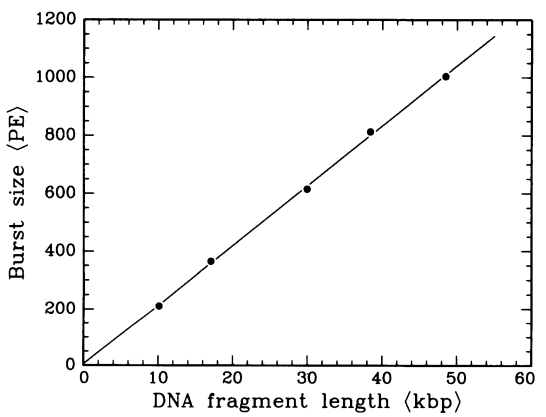

Figure 2. Correlation of the centroids of the burst sizes obtained from the fit with known DNA fragment lengths. The correlation coefficient for the linear regression fit is R = 0.9996. The fitted line passes through the origin within the precision of the fit.

Future work will concentrate on: improving the resolution of the measurement; analyzing larger fragments; and sorting individual fragments based on the size measurement. The resolution of the measurements can be improved by increasing the collection efficiency, by assuring that each DNA fragment receives the same irradiance while it is in the probe volume formed by the intersection of the laser beam and the optical collection volume, and by assuring that the efficiency of the light collection from the probe volume is uniform. In principle, the CV can approach $\sqrt{n}/n$ where n is the smaller of the number of photoelectrons collected from each fragment or the number of tags on each fragment. In practice, the CV in flow cytometric analysis is often limited by instrument related effects. Coefficients of variation in flow cytometry analyses have been demonstrated to be as low as 1%. For example, Bartholdi et al. reported CVs of about 1.5% for fluorescently stained chromosomes (15). Measured CVs as low as 0.7% have been observed routinely for 2 μm diameter Polyscience microspheres of the type described above (16).

The size of the fragments that can be analyzed by this technique is limited by the stability of the fragment against breakage during handling and analysis. In contrast to gel electrophoresis analysis, the resolution in our approach remains constant or improves as the fragment size gets larger because the fluorescence intensity from each fragment increases, decreasing the relative amount of shot noise. Furthermore, in gel electrophoresis, migration of fragments is nonlinear with fragment size (12, 17), leading to degradation of resolution as fragment size increases. By contrast, our instrument response is linear with fragment length; therefore, the potential exists for resolution improvement for larger fragments. Fragments of free DNA in excess of 250,000 bp are handled routinely in other applications (18).

We have considerable expertise in the use of flow cytometers to sort biological objects by their fluorescence properties. Los Alamos National Laboratory houses the National Flow Cytometry Resource (19). We intend to exploit the sorting capabilities of flow cytometers to sort individual DNA fragments by size. Some developmental work will be necessary to accomplish this task because our flow velocities are considerably slower than in conventional flow cytometers.

The fact that these measurements were made with 30 mW of power from a cw laser means that a relatively simple apparatus

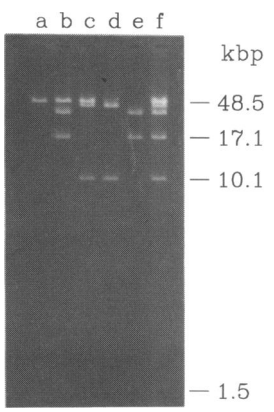

Figure 3. Gel electrophoresis of λ DNA, a *Kpn* I digest of λ DNA, and an *Apa* I digest of λ DNA, as described in Materials and Methods. Fragment length in kbp is shown at the right. Lanes were loaded with the following: (a) λ DNA; (b) λ DNA + *Kpn* I digest; (c) λ DNA + *Apa* I digest; (d) *Apa* I digest only; (e) *Kpn* I digest only; (f) λ DNA + *Kpn* I digest + *Apa* I digest.

could be constructed using an inexpensive, air-cooled Ar^+ laser and a photomultiplier tube operating in the current mode to handle the stronger signals expected from larger fragments of DNA.

Many applications that currently use gel electrophoresis to size and separate DNA fragments are foreseen for this rapid and relatively simple technique. They include new, gel-based approaches to DNA sequencing (20, 21) and DNA fingerprinting for medical diagnostics and forensic studies (22). We expect additional applications which take advantage of the speed, small sample size requirements, and accuracy of this approach will be developed.

ACKNOWLEDGMENTS

This work is supported by the Los Alamos Center for Human Genome Studies under United States Department of Energy contract W-7405-ENG-36 and the National Flow Cytometry Resource, NIH RR 01315, and is performed under the auspices of US DOE Cooperative Research and Development Agreement No. LANL-C-91-001. The authors benefited greatly from useful discussions with Mark Hammond, presently at Campbell University, Department of Biology, Buies Creek, NC, and technical assistance from Harvey Nutter and Jon Longmire, Los Alamos National Laboratory. Supported by the U.S. Department of Energy and the National Flow Cytometry Resource (NIH Grant RR 01315).

REFERENCES

1. Ogden, R. C. and Adams, D. A. (1987) In Berger, S. L. and Kimmel, A. R. (eds.), Methods in Enzymology. Guide to Molecular Cloning Techniques. Academic Press, New York, Vol. 152, pp 61–87.
2. Schwartz, D. C. and Cantor, C. R. (1984) *Cell*, **37**, 67–75.
3. Swerdlow, H., Zhang, J. Z., Chen, D. Y., Harke, H. R., Grey, R., Wu, S., Dovichi, N. J. and Fuller, C. (1991) *Anal. Chem.*, **63**, 2835–2841.
4. Kostichka, A. J., Marchbanks, M. L., Brumley, R. L., Jr., Drossman, H. and Smith, L. M. (1992) *Bio/Technology*, **10**, 78–81.
5. Hammond, M. L., Jett, J. H., Keller, R. A., Marrone, B. L. and Martin, J. C. (1992) U.S. Patent 75,091, pending.
6. Shera, E. B., Seitzinger, N. K., Davis, L. M., Keller, R. A. and Soper, S. A. (1990) *Chem. Phys. Lett.*, **174**, 553–557.
7. Wilkerson, C. W., Jr., Goodwin, P. M., Ambrose, W. P., Martin, J. C. and Keller, R. A. (1993) *Appl. Phys. Lett.*, submitted.
8. Rye, H. S., Yue, S., Wemmer, D. E., Quesada, M. A., Haugland, R. P., Mathies, R. A. and Glazer, A. N. (1992) *Nucleic Acids Res.*, **20**, 2803–2812.
9. Peck, K., Stryer, L., Glazer, A. N. and Mathies, R. A. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 4087–4091.
10. The coefficient of variation (CV) is defined to be the standard deviation of a distribution divided by its mean. The values listed in Table 1 were taken from the Gaussian fits to the data.
11. Holm, D. M. and Cram, L. S. (1973) *Exp. Cell Res.*, **80**, 105–110.
12. Reese, H. R. and Zimm, B. H. (1990) *J. Chem. Phys.*, **92**, 2650–2662.
13. The deviations from of the data from the linear regression fit shown in Figure 2 are defined by:
δ = (fitted burst size−measured burst size)/fitted burst size.
14. Johnson, M. E., Goodwin, P. M., Ambrose, W. P., Martin, J. C., Marrone, B. L., Jett, J. H. and Keller, R. A. (1993) Proceedings SPIE. SPIE, Bellingham, WA, Vol. 1895, in press.
15. Bartholdi, M. F., Sinclair, D. C. and Cram, L. S. (1983) *Cytometry*, **3**, 395–401.
16. Martin, J. C., Los Alamos National Laboratory, unpublished results.
17. Southern, E. M. (1979) *Anal. Biochem.*, **100**, 319–323.
18. Reynolds, R. J. and Lam, L. H. (1988) In Friedberg, E. C. and Hanawalt, P. C. (eds.), Handbook of DNA Repair Techniques. Marcel Dekker, Inc., New York, Vol. III, pp 427–442.
19. Sklar, L. A. (1991) *Laser Focus World*, **27**, 61–81.
20. Strathmann, M., Hamilton, B. A., Mayeda, C. A., Simon, M. I., Meyerowitz, E. M. and Palazzolo, M. J. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 1247–1250.
21. Studier, F. W. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 6917–6921.
22. Brettell, T. A. and Saferstein, R. (1991) *Anal. Chem.*, **63**, 148R–164R.