Capillary gel electrophoresis for rapid, high resolution DNA sequencing

Harold Swerdlow and Raymond Gesteland
Departments of Human Genetics and Bioengineering, Howard Hughes Medical Institute, University of Utah, Salt Lake City, UT 84132, USA

Received December 27, 1989; Revised and Accepted February 20, 1990

ABSTRACT

Capillary gel electrophoresis has been demonstrated for the separation and detection of DNA sequencing samples. Enzymatic dideoxy nucleotide chain termination was employed, using fluorescently tagged oligonucleotide primers and laser based on-column detection (limit of detection is 6,000 molecules per peak). Capillary gel separations were shown to be three times faster, with better resolution (2.4 ×), and higher separation efficiency (5.4 ×) than a conventional automated slab gel DNA sequencing instrument. Agreement of measured values for velocity, resolution and separation efficiency with theory, predicts further improvements will result from increased electric field strengths (higher voltages and shorter capillaries). Advantages of capillary gel electrophoresis for automatic DNA sequencing instruments and for genomic sequencing are discussed.

INTRODUCTION

DNA sequencing technology, both conventional and automated, has classically been constrained by the inability of slab gel electrophoresis to resolve fragments representing more than 600 consecutive nucleotides in a single run[1-5]. To transcend this limitation, and improve the speed of automated DNA sequencing, capillary electrophoresis technology appears to be an attractive alternative.

Capillary electrophoresis is a developing analytical and micro-preparative tool[6-9]. The technique provides faster separations, at higher resolution and with greater separation efficiencies than conventional electrophoresis[10,11]. These improvements are a consequence of the high voltages (10−30 kV typically) that can be utilized in fused silica micro-capillaries (10−100 μm inner diameter) because of efficient heat dissipation[8]. For many applications, there is no need for an anti-convective medium (e.g. agarose or polyacrylamide) due to the dominant frictional forces at the capillary wall. Impressive separations have been obtained in capillary zone electrophoresis without any supporting medium other than the capillary itself[12,13]. However, fractionation of polymers, e.g. sodium dodecyl sulfate treated proteins and denatured nucleic acids, often requires a gel matrix for sieving[14,15].

Karger has shown that polyacrylamide gel filled capillaries can be used to achieve single nucleotide separations of DNA oligonucleotide markers, 40−60 nucleotides in length, using UV absorbance detection[15]. In the present work, we demonstrate the feasibility of sequencing DNA in polyacrylamide gel filled capillaries, utilizing ultra-sensitive fluorescence detection[16,17], and actual DNA sequencing samples.

A direct comparison of automated slab-gel and capillary techniques is shown, verifying enhanced speed, resolution and efficiency, and testing the higher performance against predictions made by electrophoretic theory[8].

EXPERIMENTAL

Apparatus

The capillary electrophoresis system (Fig. 1) is assembled on a 3×4 foot optical table using standard optical mounts and experiments are performed in a dark room. Samples were introduced electrophoretically at one end of the capillary using a 2 kV power supply (EC Corporation model 500). For electrophoresis, this end of the capillary was then placed in an insulated, safety-interlocked cathode buffer chamber (at negative potential with respect to the anode chamber and optical table − earth ground). The −30 kV power supply (Glassman High Voltage model mj30N0400) was operated through a custom-made microprocessor-based controller board. This board constantly monitored safety interlocks, controlled voltage or power, and performed volt-hour integration. DNA was detected by laser-induced fluorescence. The 488 nm, .8 mm diameter beam (Ion Laser Technology 100 mW air-cooled argon-ion laser model 550ASL operating at 30 mW single line, continuous wave) was filtered spectrally with a laser line filter (488 nm center wavelength, 10 nm bandwidth bandpass interference filter - Microcoatings ML1-488) and then spatially with a 1.5 mm diameter pinhole. The beam was focused at the center of the capillary to an approximately 20 μm dia. spot size by a 7.5 mm dia. 18 mm focal length plano-convex lens. This produced an effective detection volume of approximately 3×10^{-8} ml (30 picoliters). In the detection region, the polyimide buffer coating on a 1 cm long section of capillary was removed by a fine bunsen-burner flame. The detection region was surrounded by a quartz fluorometer cuvette to reduce dust contamination. The laser beam travelled parallel to the table surface, and the capillary was oriented vertically, but tilted at an angle of 25 degrees towards the laser. This tilt angle prevented the majority of scattered laser light from reaching the photomultiplier tube. The light was

collected with a .40 numerical aperture 20× microscope objective (Swift), oriented horizontally, and at right angles to the laser beam. The objective focused the central illuminated region onto a rectangular aperture (spatial filter) to reduce scattered light which originated from the capillary walls. The light then passed through two colored glass filters (Schott 3 mm thick OG-515) and a bandpass interference filter (Omega Optical model 530DF30; 530 nm center wavelength, 30 nm bandwidth), before entering a light tight box containing the photomultiplier tube (Hamamatsu R1527 low-noise bialkali, operated at 700 V). The photomultiplier tube current was converted to a voltage signal (100 V/μA) by a low-noise operational amplifier circuit. The voltage output was filtered through an analog 1 Hz low pass filter before entering the analog to digital converter (Metrabyte Corp. DAS-8PGA). The acquisition (at 2 data points per second), storage, and display of data was accomplished on an 80286-based personal computer.

MATERIALS AND METHODS

Preparation of capillary gels

A 60 cm section of capillary (75 μm i.d., 375 μm o.d. fused silica – Polymicro Technologies) was heated in a fine bunsen-burner flame to remove all traces of the polyimide buffer coating in a 1 cm region near its midpoint. A 5.7% acrylamide, .3% bis-acrylamide (6%T, 5%C – both BioRad electrophoresis purity reagents), 8M urea (Schwarz/Mann ultra pure), 1× TBE (89 mM Tris – BRL ultra pure enzyme grade, 89mM Boric acid – CMS chempure, 2 mM EDTA – Sigma) solution was filtered through a .45 μm filter and then degassed under vacuum (aspirator) at 50°C in an ultrasonic bath for 5 min. Immediately after degassing, the volume was adjusted to its original level with degassed, filtered water, then the solution was brought to .03% v/v TEMED and .03% w/v ammonium persulfate (both were BioRad electrophoresis purity; ammonium persulfate 10% solution was freshly prepared; note, these concentrations are lower than normally used for sequencing gels[18]). One end of the capillary was inserted into a reservoir chamber which was elevated 10 cm and the acrylamide solution was added. After polymerization for 2 hr, gels were examined under a microscope for bubbles, which occasionally formed. The ends of the capillary were trimmed to 50 cm long; the detector region for the run displayed in Fig. 2 was 29 cm from the sample end.

Preparation of samples

Single stranded m13mp18 DNA containing a salmonella hisR tRNA gene insertion (a gift of Norma Wills) was used as a template in the DNA sequencing reaction. The DNA template (2 pmol) and fluorescently labelled M-13 universal sequencing primer[19] (4 pmol -Applied Biosystems –21 'FAM' primer cat # 400836) were annealed in the dark by heating to 65°C in 100 μl of sequenase buffer (10 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 50mM NaCl) for 5 min and then allowed to cool to 37°C slowly for 1 hr. Dideoxy CTP mix (final concentrations: 200 μM dATP, dTTP, dCTP, 300 μM 7-deazo GTP, 5 μM dideoxy CTP–Pharmacia) and Sequenase (40 units–USB) were added to the DNA in sequenase buffer (plus 7 mM dithiothreitol) in a final volume of 150 μl. Reactions were incubated, in the dark, 10 min at 37°C then 10 min at 70°C to inactivate the sequenase. Reaction mixtures were then ethanol precipitated and resuspended in 2 μl of loading buffer (10 mM EDTA pH 8.0 in de-ionized formamide).

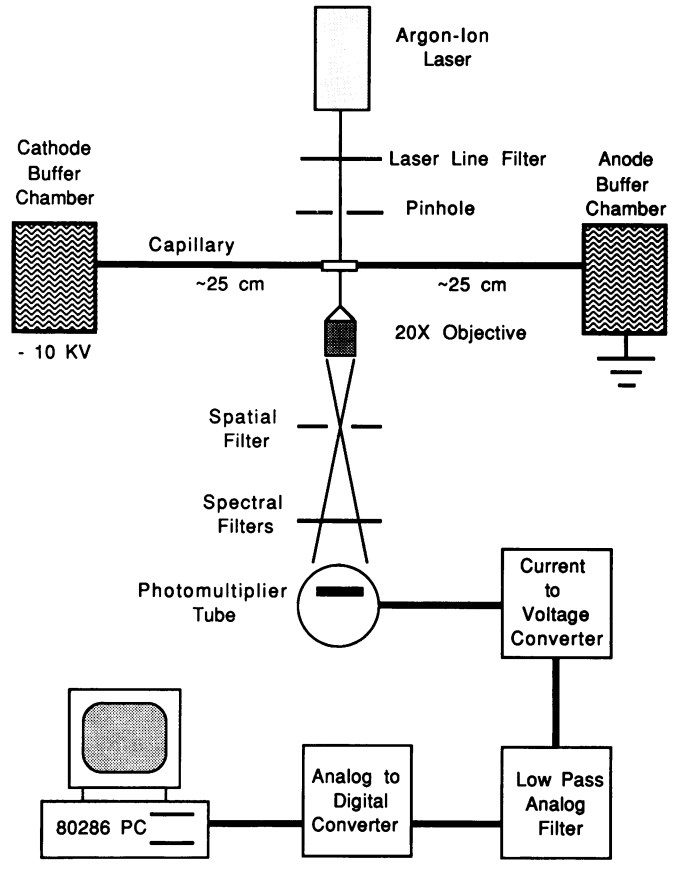

FIGURE 1. A schematic illustration of the capillary gel electrophoresis apparatus. A complete description of the system components is contained in Materials and Methods.

Gel Electrophoresis

Capillary gels were pre-run at 9 kV in 1× TBE buffer for 60 min. Samples were heated at 95°C for 1 min in the dark and introduced electrophoretically (5 sec at 2 kV) into the capillary. Assuming no sample concentration effects, the injected volume is calculated using column geometry, the injection voltage and time, and the running voltage and retention time of a peak. Using the C100 peak, e.g., the amount of sample loaded on the gel was .4 nl. Gels were run at 9 kV (180 V/cm, 5–10 μA) at ambient temperature.

The slab gel in Fig. 3b was a 6% polyacrylamide sequencing gel formed and handled according to Applied Biosystems protocols[19]. The gel was 40 cm long, with a sample to detector length of 25 cm. 6 μl of a 1:25 dilution of the same sample used in Fig 2a was loaded in a single lane. Run at a constant power of 35 Watts (approx. 45°C), the gel required an average of 1400 volts. 6000 data points were collected for 630 min beginning 80 min after sample loading.

RESULTS AND DISCUSSION

In order to assess the potential of capillary gel electrophoresis for DNA sequencing, we constructed a prototype electrophoresis apparatus in conjunction with a simple, high sensitivity, on-column, laser-based fluorescence detector (Fig. 1). A 50 cm long fused silica capillary (75 μm inner diameter) was passively filled

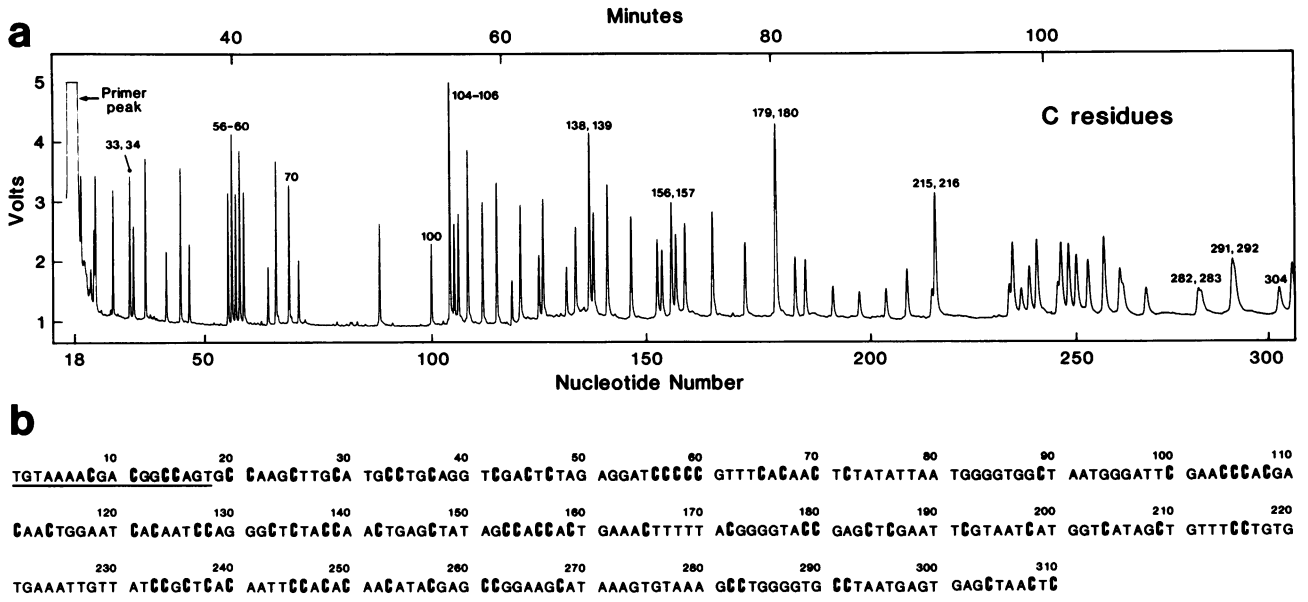

FIGURE 2. Capillary gel electrophoresis of a DNA sequencing reaction performed with a fluorescently tagged primer and terminated with dideoxy CTP. a) The raw digitized data, obtained by monitoring fluorescence at a fixed point. The abscissa represents time after sample introduction, and is displayed in minutes as well as corresponding oligonucleotide length; fluorescent signal is plotted on the ordinate as voltage. b) The corresponding known DNA sequence of the chain-extended strand; numbering begins with the first nucleotide of the 18-mer oligonucleotide primer (underlined). Each peak in a) maps to a C residue (bold type) in b). Assignment of nucleotide numbers to peaks was assisted by comparison of the pattern with a conventional sequencing gel autoradiogram of the same sequence. The absence of doublet peaks corresponding to C179-C180 and to C291-C292, is probably due to the capillary gels being run at room temperature, stabilizing secondary structure in these regions.

with a 6% acrylamide/bis-acrylamide monomer solution (6%T, 5%C)[20] and the gel polymerized *in situ*. An actual sequencing sample was used to test the system. This sample was prepared by chain-extending a fluorescein labelled 18 nucleotide synthetic primer on a complementary template with DNA polymerase and deoxynucleoside triphosphates in the presence of dideoxy cytidine triphosphate. Sample, containing the C-terminated fragments, was introduced electrophoretically into the capillary gel. Electrophoresis was performed at 9000 V (180 V/cm), moving the oligonucleotides past the fixed detector, 29 cm from the sample end. The resulting raw data are displayed in Fig. 2a, with fluorescent signal plotted against time. Each peak in Fig. 2a corresponds to a C residue (bold type) in the known sequence, Fig. 2b. Baseline resolution of peaks differing by one residue is evident through at least 160 nucleotides (C56-C60, C156-C157), and doublet C peaks (C282-C283) can be identified out to 300 bases.

In order to directly compare capillary gel and conventional slab-gel techniques, identical samples were analyzed on our instrument and on a commercial fluorescent detection based machine (Applied Biosystems model 370A) in Fig. 3. The capillary electrophoresis was more than 3 times faster, taking 120 min to reach nucleotide C304, while the conventional gel required 370 min. Resolution can be calculated using the IUPAC formula,

$$R = \Delta T/4\sigma_T$$

where ΔT is the difference in time of elution between two consecutive C peaks (differing by one nucleotide) and σ_T is the standard width of a single peak[21]. Resolution values obtained from equation 1 for the capillary range from 2.3 for the C33-C34 doublet, 1.8 for C156-C157, and .8 for C215-C216. These values compare with .85, .6, and .4 for the same pairs, measured on the automated system. A standard indicator of separation efficiency, borrowed from chromatography, is the number of theoretical plates N, given by:

$$N = (T/\sigma_T)^2 \qquad (2)$$

where T is the total time for a given species to elute[21]. For the data in Fig. 2a, using equation 2, a maximum value of 2.9 million theoretical plates was found for residue C100 (10 million theoretical plates per meter of separation). This high efficiency compares favorably with a value of 650,000 plates (5 million/meter) reported previously by Karger for capillary gel electrophoresis[15], but recently, far higher values have been obtained by his group (15 million plates or 30 million/meter[22]). The efficiency measured for the automated slab-gel separation of Fig. 3b is lower throughout, e.g. 460,000 plates (1.8 million/meter) for C100. By both estimates, the capillary method gave superior performance.

Zone electrophoresis theory makes specific predictions for the relative performance of the capillary and slab-gel techniques[8]. The standard equations need to be modified due to the fact that the sample to detector distance l in our apparatus is not the same as the overall gel length L. The velocity v of migration of a single species of electrophoretic mobility μ under the influence of a voltage V is given by:

$$v = \mu V/L. \qquad (3)$$

The resolution of two species possessing an average electrophoretic mobility μ_{ave}, a difference in mobility $\Delta\mu$, and a diffusion coefficient D, is given by:

$$R = .177\Delta\mu(Vl/D\mu_{ave}L)^{1/2},$$

where the number of theoretical plates is given by:

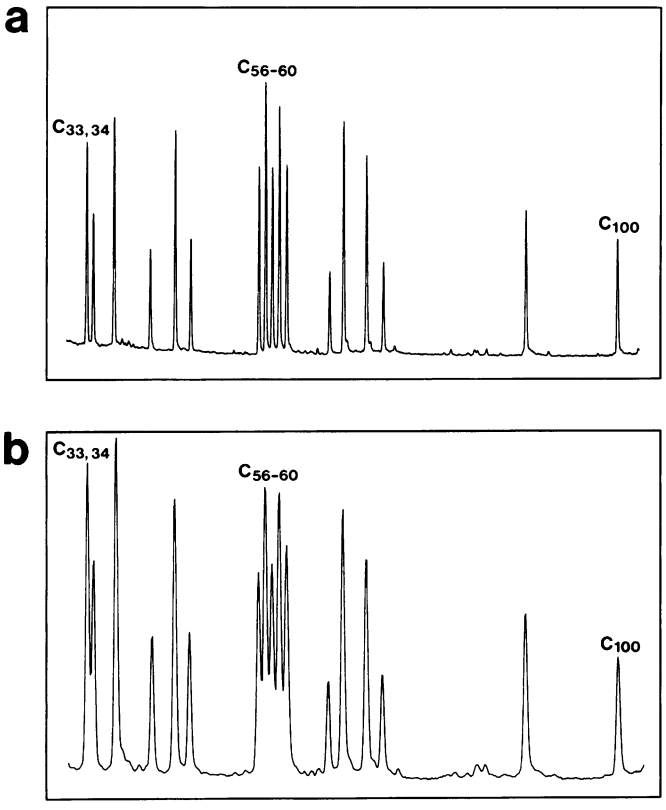

FIGURE 3. Comparison of capillary and slab-gel electrophoresis of a DNA sequencing reaction. a) a portion of Fig. 2a encompassing nucleotides C33 to C100. b) The raw data from an Applied Biosystems model 370A DNA sequencing instrument used to run the identical sample. The peak labelled C100 took 54 min to reach the detector in a) and 150 min in b).

$$N = \mu Vl/2DL. \quad (5)$$

Forming a ratio of the values predicted for each system, eliminates the mobility and diffusion terms for individual species, and predicts that for equal geometries, the velocity and plate count are proportional to voltage while resolution is proportional to the square root of voltage. Using the run parameters of Figs. 2a and 3b, i.e. L= 50 cm, l= 29 cm and V = 9000 Volts for the capillary, and L= 40 cm, l= 29 cm and V = 1400 Volts for the slab gel, the predicted ratios (capillary/slab gel) for velocity, resolution and theoretical plates from equations 3−5 were 5.1, 2.4, and 6.0 respectively. The measured ratios (averaged over at least 5 peaks/doublets) were 3.4, 2.4, and 5.4 respectively. The theory agreed very well with the experimental data for both resolution and plate counts, and predicts that further improvements will result from increased voltages. The velocity for capillary gel electrophoresis was lower than that predicted by theory. This discrepancy is likely to be associated with the difference in running temperature of the two systems (near room temperature for the capillary and approx. 45°C for the slab gel)[23,24]. An additional 33% reduction in analysis time is thus probable if the gels are run near 50°C (this should also reduce sequencing artifacts − see Fig. 2).

Capillary gel DNA sequencing can take advantage of small sample sizes (sub nanoliter) provided detectors of adequate sensitivity are used. Our relatively simple, on-column laser fluorescence detector can detect a 10^{-11} M (10 picomolar) solution of fluorescein flowing continuously in an open capillary (signal to noise ratio of 3:1). This limit of detection corresponds to 3×10^{-22} mol, 10^{-19} g or less than 200 molecules of dye present in the detected volume. An average peak (e.g. C100 in Fig. 2a) contains approximately 130,000 molecules, 10^{-14} g or only .2 attomoles (0.2×10^{-18} mol) of DNA. Using the measured baseline noise standard deviation and the height of this C100 peak, the smallest discernible DNA sequencing band in our capillary gel system can be estimated to be .01 amol, or 6,000 molecules. By comparison, the quantity of DNA in a conventional radiolabelled sequencing gel band is 100 to 1000 amol[25], while the detection limits quoted for fluorescence based sequencing instruments vary from 1−10 amol to 50 amol of DNA per band[2,4,5,25]. In capillaries, Dovichi has achieved the lowest reported mass sensitivity limits by employing a sheath-flow cuvette based detector (.0017 amol of fluorescently tagged amino acids in a peak)[26]. We are currently working with Dr. Dovichi to obtain further improvements in sensitivity and resolution, with simultaneous detection and discrimination of all four base specific reactions[1,4].

CONCLUSION

Certain problems still need to be addressed before capillary gel electrophoresis can become the method of choice for DNA sequencing. At high voltages, electroendosmosis, caused by the presence of ionized silanol groups on the capillary's inner surface, can cause gels to migrate out of the tubes[14]. This limits the voltages that can be utilized for enhanced performance and increased speed. Air bubble formation in the capillary, during polymerization, occurs about 20% of the time. Bubbles which arise during electrophoresis, remain a problem. These latter bubbles form near the injection end of the capillary and often seriously impair resolution, especially of larger fragments. Cutting a few centimeters of the capillary at the injection end after the sample has passed, usually relieves the problem temporarily. It is expected that commerical production of quality-controlled gel filled capillaries may alleviate some of the problems of pouring and running capillary gels[2]2.

Major improvements in DNA sequencing gels have been minimal in the many years since their inception[27]. The success of genomic sequencing efforts relies on significant increases in DNA sequencing throughput and cost effectiveness. Current automated DNA sequencing practice has been constrained by slow speed and the inability of slab-gel electrophoresis to reliably fractionate molecules larger than 600 nucleotides in length. As a result of higher voltages, capillary gel electrophoresis, with its improved speed and resolution, may provide a method to significantly improve the throughput. The importance of extending the length of a fragment that can be sequenced in a single run is two fold. First, fewer clones and fewer nucleotides need to be sequenced, but additionally, larger fragments result in a much easier task of overlapping sequences[28].

The cost benefits arising from reduction in sample and gel reagent requirements in capillary methods, may be substantial. Additionally, major savings could result from the reduction of labor costs associated with an automated capillary technology. Commercial capillary electrophoresis instruments are available, capable of multiple sample loadings and automated runs[9,29]. Commerical availability of reliable gel filled capillaries is likely in the near future. Ultimately, when when capillary sequencing gels can be multiplexed, i.e. electrophoresed and analyzed in parallel, large-scale DNA sequencing endeavors, such as the

Human Genome Project, may take advantage of this emerging technology.

ACKNOWLEDGEMENTS

The authors would like to thank Jeff Ives, Robert Weiss and Norman J. Dovichi for substantial technical advice and for critical reading of the manuscript, Debra Milton, Joel Harris and Edward S. Yeung for scientific consultations, Norma Wills and Diane Dunn for supplying reagents, and especially Margaret Robertson for consultations, various components, and for performing the separation in Fig. 3b. This work was supported by grants from the National Science Foundation to H.S. and from the Department of Energy to R.G.

REFERENCES

1. Connell, C., Fung, S., Heiner, C., Bridgham, J., Chakerian, V., Heron, E., Jones, B., Menchen, S., Mordan, W., Raff, M., Recknor, M., Smith, L., Springer, J., Woo, S. and Hunkapiller, M. (1987) *Biotechniques* **5**, 342–348.
2. Ansorge, W., Sproat, B., Stegemann, J., Schwager, C. and Zenke, M. (1987) *Nucleic Acids Res.* **15**, 4593–4602.
3. Toneguzzo, F., Beck, J., Cahill, P., Ciarkowski, M., Page, G., Glynn, S., Hungerman, E., Levi, E., Ikeda, R., McKenney, K., Schmidt, P. and Danby, P. (1989) *Biotechniques* **7**, 866–877.
4. Prober, J.M., Trainor, G.L., Dam, R.J., Hobbs, F.W., Robertson, C.W., Zagursky, R.J., Cocuzza, A.J., Jensen, M.A. and Baumeister, K. (1987) *Science* **238**, 336–341.
5. Brumbaugh, J.A., Middendorf, L.R., Grone, D.L. and Ruth, J.L. (1988) *Proc. natn. Acad. Sci. U.S.A.* **85**, 5610–5614.
6. Virtanen, R. (1974) *Acta Polytech. Scand.* **123**, 1–67 .
7. Mikkers, F.E.P., Everaerts, F.M. and Verheggen, TH.P.E.M. (1979) *J. Chromatogr.* **169**, 11–20.
8. Jorgensen, J. and Lukacs, K.D. (1981) *Analyt. Chem.* **53**, 1298–1302.
9. Karger, B.L. (1989) *Nature* **339**, 641–642.
10. Ewing, A.G., Wallingford, R.A. and Olefirowicz, T.M. (1989) *Analyt. Chem.* **61**, 292A–303A.
11. Gordon, M.J., Huang, X., Pentoney, S.L.Jr. and Zare, R.N. (1988) *Science* **242**, 224–228.
12. Lauer, H.H. and McManigill, D. (1986) *Analyt. Chem.* **58**, 166–170.
13. Cohen, A. S., Najarian, D., Smith, J.A. and Karger, B.L. (1988) *J. Chromatogr.* **458**, 323–333.
14. Cohen, A.S. and Karger, B.L. (1987) *J. Chromatogr.* **397**, 409–417.
15. Cohen A.S., Najarian, D.R., Paulus, A., Guttman, A., Smith, J.A. and Karger, B.L. (1988) *Proc. natn. Acad. Sci. U.S.A.* **85**, 9660–9663.
16. Kuhr, W.G. and Yeung, E.S. (1988) *Analyt. Chem.* **60**, 2642–2646.
17. Cheng, Y.F. and Dovichi, N.J. (1988) *Science* **242**, 562–564.
18. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
19. *User Bulletin* #7, Applied Biosystems Model 370A DNA Sequencer. Applied Biosystems, Foster City, CA.
20. Hjerten, S. (1962) *Arch. Biochem. Biophys. Suppl.* **1**, 147–151.
21. Giddings, J.C. (1969) *Sep. Sci.* **4**, 181–189.
22. Guttman, A., Cohen, A.S., Heiger, D.N. and Karger, B.L. (1990) *Analyt. Chem.* **62**, 137–141.
23. Grushka, E., McCormick, R.M. and Kirkland, J.J. (1989) *Analyt. Chem.* **61**, 241–246.
24. Wieme, R.J. (1975) In Heftmann, E. (ed.), *Chromatography: A Laboratory Handbook of Chromatographic and Electrophoretic Methods*. 3rd ed., Van Nostrand Reinhold, New York, Chapter 10.
25. Smith, L.M., Sanders, J.Z., Kaiser, R.J., Hughes, P., Dodd, C., Connell, C.R., Heiner, C., Kent, S.B.H. and Hood, L.E. (1986) *Nature* **321**, 674–679.
26. Wu, S. and Dovichi, N.J. (1989) *J. Chromatogr.* **480**, 141–155.
27. Maxam, A.M. and Gilbert, W. (1977) *Proc. natn. Acad. Sci. U.S.A.* **74**, 560–564.
28. Lander, E.S.; Waterman, M.S. (1988) *Genomics* **2**, 231–239.
29. Compton, S.W. and Brownlee, R.G. (1988) *Biotechniques* **6**, 432–440.