A System for Rapid DNA Sequencing with Fluorescent Chain-Terminating Dideoxynucleotides

JAMES M. PROBER,* GEORGE L. TRAINOR, RUDY J. DAM, FRANK W. HOBBS, CHARLES W. ROBERTSON, ROBERT J. ZAGURSKY, ANTHONY J. COCUZZA, MARK A. JENSEN, KIRK BAUMEISTER

A DNA sequencing system based on the use of a novel set of four chain-terminating dideoxynucleotides, each carrying a different chemically tuned succinylfluorescein dye distinguished by its fluorescent emission is described. Avian myeloblastosis virus reverse transcriptase is used in a modified dideoxy DNA sequencing protocol to produce a complete set of fluorescence-tagged fragments in one reaction mixture. These DNA fragments are resolved by polyacrylamide gel electrophoresis in one sequencing lane and are identified by a fluorescence detection system specifically matched to the emission characteristics of this dye set. A scanning system allows multiple samples to be run simultaneously and computer-based automatic base sequence identifications to be made. The sequence analysis of M13 phage DNA made with this system is described.

THE DEVELOPMENT OF SEQUENCING METHODS FOR DETERmining the order of nucleotide bases in deoxyribonucleic acid (DNA) has led to rapid advances in our understanding of the organization and processing of information in biological systems. Two methods are available for DNA sequencing: the chemical degradation method of Maxam and Gilbert (*1*), and the dideoxy chain termination method of Sanger (*2*). Traditionally, each approach affords four sets of radioisotopically labeled fragments which are resolved according to their lengths by gel electrophoresis and the resulting autoradiographic pattern is used to obtain the DNA sequence.

The Maxam-Gilbert and Sanger techniques, which are conceptually elegant and efficacious, are in practice time-consuming and labor-intensive, partly because a single radioisotopic reporter is used for detection. Using one reporter to analyze each of the four bases requires four separate reactions and four gel lanes. The resulting autoradiographic patterns, obtained after a delay for exposure and development, are complex and require skilled interpretation and data transcription.

These deficiencies can be corrected by switching from a radioisotopic to a fluorescent reporter. We describe here a system for DNA sequencing in which four chemically related, yet distinguishable, fluorescence-tagged dideoxynucleotides are used to label DNA by a modified Sanger protocol with a suitable chain-extending DNA polymerase. The fluorescent sequencing fragments are resolved temporally rather than spatially in a single lane by conventional polyacrylamide gel electrophoresis. Analysis of the fluorescent emission of each fragment permits us to identify the terminating nucleotide and assign the sequence directly in real time.

Fluorescent tags. We have developed a family of fluorescent dyes with largely overlapping yet distinct emission bands. These dyes are 9-(carboxyethyl)-3-hydroxy-6-oxo-6H-xanthenes or succinylfluoresceins (SF-xxx, where xxx refers to the emission maximum in nanometers) (Fig. 1A). They are readily prepared from succinic anhydride and an appropriately substituted resorcinol by a modification of the procedure cited for the parent dye, SF-505 (*3*). The fluorescent forms of these dyes are the dianions, which predominate in aqueous solution above *p*H 7. The dianion of SF-505 absorbs maximally at 486 nm, a wavelength that is well suited for excitation by an argon ion laser operating at 488 nm. This species shows an absorption coefficient of $72{,}600 M^{-1}$ cm^{-1} at the maximum and fluoresces with an emission maximum at 505 nm with a quantum yield comparable to that of fluorescein. The carboxylic acid functionality is not essential for fluorescence and is used here for covalent attachment to the nucleotides by means of standard methodologies.

The wavelengths of the absorption and emission maxima in the succinylfluorescein system are tuned by changing the substituents $-R_1$ and $-R_2$. The absorption spectra for the four dyes where $-R_1$ and $-R_2$ are $-H$ or $-Me$ are shown in Fig. 1B. Close spacing of the absorption maxima results in efficient excitation of all dyes by the single emission line of the argon ion laser. (The absorption coefficients at 488 nm of the two most disparate dyes, SF-505 and SF-526, differ only by a factor of 2.) These succinylfluoresceins carry the same charge and are nearly identical in size, minimizing any differential perturbation of the electrophoretic mobilities of the DNA fragments to which they are attached. We observe no differences in the relative mobilities among identical DNA fragments tagged with any of the four dyes.

Labeling strategy. Dideoxy sequencing involves the template-directed, enzymatic extension of a short oligonucleotide primer in the presence of chain-terminating dideoxyribonucleotide triphosphates (ddNTP's). A nested set of DNA fragments is produced

J. M. Prober, R. J. Dam, and C. W. Robertson are with the Engineering Physics Laboratory; G. L. Trainor, F. W. Hobbs, R. J. Zagursky, A. J. Cocuzza, and K. Baumeister are with the Central Research and Development Department (contribution number 4543); and M. A. Jensen is with the Medical Products Department, all of the E. I. du Pont de Nemours & Company (Inc.), Wilmington, DE 19898.
*To whom correspondence should be addressed.

having primer-defined 5′ ends and variable 3′ ends determined by the positions of incorporation of a given base (as ddNMP). Radioisotopic label is generally incorporated by including [α-^{32}P]dNTP or [α-^{35}S]dNTP in the reaction mixture such that internal nucleotides in the newly synthesized portions of the fragments are labeled.

The analogous incorporation of fluorescence-tagged nucleotides may be difficult to achieve enzymatically, and may adversely affect the crucial relation between chain length and electrophoretic mobility, which is essential for accurate sequence determination. We have chosen instead to incorporate a single fluorescent tag by labeling chain-terminating dideoxynucleotide triphosphates. This labeling approach offers several advantages. Generation of all four sets of DNA sequencing fragments can be carried out simultaneously in a single reaction since only the terminating nucleotide carries the tag. In addition, many polymerase pausing artifacts are eliminated since only those fragments resulting from bona fide termination events carry a fluorescent tag.

The set of four fluorescence-tagged chain-terminating reagents we have designed and synthesized is shown in Fig. 2A. These are ddNTP's to which succinylfluorescein has been attached via a linker to the heterocyclic base. (These reagents are designated N-xxx where N refers to the ddNTP and xxx refers to the SF-xxx portion.) The linker is attached to the 5 position in the pyrimidines and to the 7 position in the 7-deazapurines. The 7-deazapurines were used to facilitate stable linker arm attachment at that site. Coupling of the dyes to the nucleotides was found to influence both the wavelengths of the absorption and emission maxima and their relative fluorescence yields. The dyes and bases were therefore paired for maximum distinguishability and for balanced net fluorescence sensitivities.

The synthetic scheme devised for these reagents is highly convergent and general. The preparation of T-526 is illustrative. 5-Iodo-2′,3′-dideoxyuridine was coupled to *N*-trifluoroacetylpropargylamine under palladium(0) catalysis in dimethylformamide (*4*). The resulting derivatized nucleoside was converted to its 5′-triphosphate (*5*) and deacylated to afford 5-(3-amino-1-propynyl)-2′,3′-dideoxyuridine triphosphate. This amine was coupled with an *O*-acetyl–protected form of the SF-526–sarcosine conjugate and deprotected to afford T-526.

The fluorescence-tagged chain-terminators are accepted as alternative substrates by avian myeloblastosis virus (AMV) reverse transcriptase with efficiencies comparable to that of the corresponding unsubstituted ddNTP's. A comparison of the sequencing ladders produced using the ddNTP's and their fluorescent counterparts is shown in Fig. 3. The fidelity of terminator incorporation is maintained. Comparison of lanes 4 and 5, containing ddCTP and C-519, respectively, shows that the fluorescence-tagged fragments run approximately 2 bases slower than their untagged counterparts. This mobility shift is consistent for all four sets of fluorescence-tagged fragments, allowing the sequence ladder to be read and qualifying this set of reagents for single-lane DNA sequencing. These terminators are also substrates for modified T7 DNA polymerase (*6*). They are not, however, substrates for the Klenow fragment of DNA polymerase I from *Escherichia coli*.

Fluorescence detection system. Labeled DNA fragments, produced by the enzymatic chain extension reactions, are separated by polyacrylamide gel electrophoresis, detected, and identified as they migrate past the fluorescence detection system illustrated in Fig. 4. High signal-to-noise ratio is achieved in this system through efficient excitation, optical filtering, and light collection.

Strong excitation is obtained with a laser source that provides most of its output energy at a single wavelength and by matching the dye absorbances to that wavelength. Fluorescence detection is enhanced approximately fourfold by depositing a mirror on the outside surface of the glass gel support plate furthest away from the laser. The excitation beam is thereby returned through the gel, in effect doubling the excitation pathlength, and the fluorescence headed away from the detectors is also returned, increasing the amount of collected light.

In addition to the desired fluorescence from the dye-labeled DNA, light emanating from the excitation region includes scattered laser radiation, Raman scattering, and fluorescence from other sources. Optical filtering and other means are used to reduce these undesired signals. Secondary laser lines are removed with a narrow-band interference filter placed at the source. The reflected component of the 488-nm laser light is eliminated by the mirror, which provides a return path for the beam out of the detectors' field of view. Scattered light is removed by a filter stack consisting of an interference filter, a fiber-optic face plate, and a colored glass absorbing filter. The interference filter is designed to strongly reject the excitation light, pass the dye emission region, and reject Raman scattering and fluorescence above 560 nm. At high angles of incidence, the filter passband is shifted toward the excitation wavelength, thus increasing background levels and associated noise.

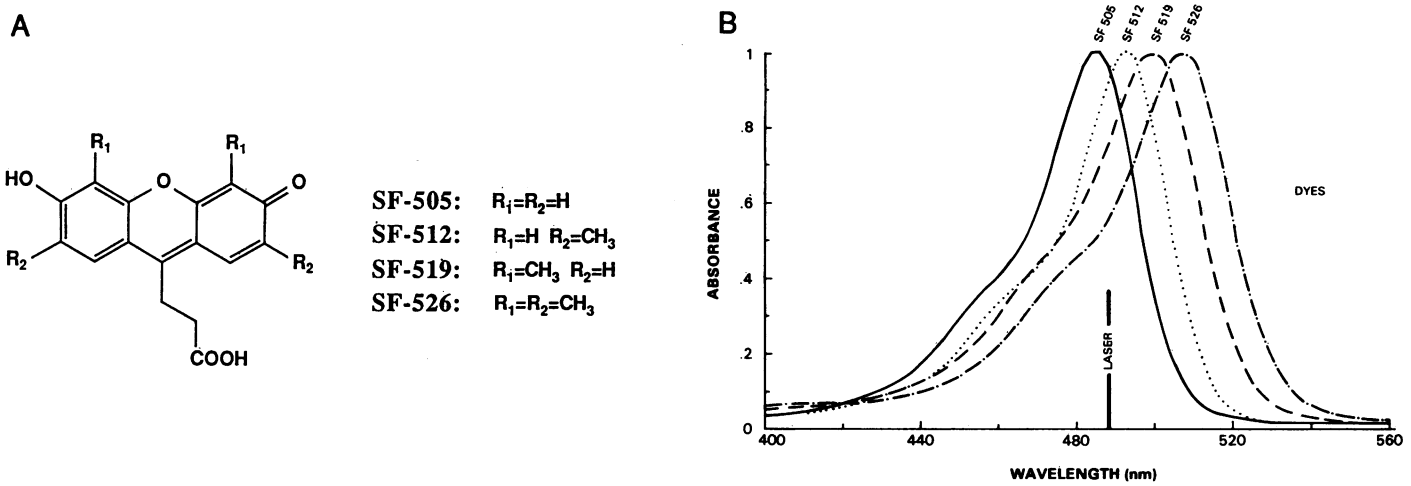

Fig. 1. Succinylfluorescein dyes. (**A**) Chemical structure of the four dyes used to label dideoxynucleotide triphosphates for use as chain-terminators in modified dideoxy DNA sequencing protocols. (**B**) Normalized absorption spectra of the dyes shown in (A). Absorption coefficient at the maximum for SF-505 is 72,600M^{-1} cm^{-1}. Spectra were measured in *p*H 8.2, 50 m*M* aqueous tris-HCl buffer. The other dyes have coefficients within 10 percent of this value. Vertical bar (laser) indicates the position of the argon ion laser line at 488 nm used for fluorescence excitation.

Thus a fiber-optic face plate with extramural absorber is employed as an aperture to restrict entrance angles on the filter. This element absorbs incident rays outside the acceptance angle of the fiber, in this case about 22 degrees. This aperture appears over any point of excitation across the gel.

The dyes that we used are distinguished by small differences in their absorption and emission spectra. Close spacing of the spectra facilitates efficient excitation and simultaneous detection. This contrasts with the use of large differences in emission spectra that one would intuitively select to facilitate discrimination. Figure 2B shows the emission spectra of the fluorescence-tagged chain-terminators that occur in a polyacrylamide gel. In our system, two photomultiplier tube detectors, each with a different filter stack, view the fluorescence simultaneously; one collects light at the low-wavelength side of the emission bands and the other collects light at the high wavelength side. When, for example, G-505–terminated fragments pass the detectors, one detector registers a large signal relative to the other; when T-526–terminated fragments pass, the reverse occurs. The A-512– and C-519–terminated fragments give specific detector signal ratios lying between these two extremes. The ratio of baseline-corrected peak intensities is used to determine the base assignment. This method provides very efficient light collection and obviates the need for more refined spectral analysis.

The scanning mechanism shown in Fig. 4 directs the laser beam to the gel through a periscope mounted on the shaft of a stepper motor. By stepping the periscope, any point in the detection region can be accessed. The detectors span the full width of a gel so that fluorescence is collected and analyzed from wherever the laser is pointed. Directing only the laser beam, rather than any of the major optical elements, results in a low-noise scanning system that is mechanically simple and stable.

Data processing and analysis. Processing of the raw detector data and subsequent nucleotide base assignment is accomplished in several steps. Digitized signals from both detectors are subjected to a peak-finding algorithm to determine the time positions of the DNA bands. Once the DNA peaks are located, the ratios of baseline-corrected peak intensities are calculated along with a statistical calculation of the certainty of those ratios. The 15 percent most certain ratios are displayed as a histogram in Fig. 5. These fall naturally into clusters centered about four values, each defining one of the bases. These reference values can differ slightly from lane to lane due to geometrical effects of the wavelength selection filters. To allow for these variations, the reference values are determined independently for each lane.

The four reference values correspond to peaks in the histogram. Measurement of scatter about the reference values allows the determination of boundaries which are used for base assignment. When DNA peaks are no longer resolvable or when their amplitude falls below the inherent noise level of the detection system, no further bases are assigned. In the current system, this usually occurs between 300 to 400 bases from the 3′ terminus of the primer. Deconvolution, predictive modeling, and digital signal processing are being investigated to extend the readable sequence range.

Nucleotide sequence determination. Raw detector data obtained from the nucleotide sequence of the phage M13*mp*18 during a 6-hour run are shown in Fig. 6. A modified, two-stage dideoxy chain termination protocol (*6*) was used to generate the DNA fragments for fluorescence-based sequence analysis. The data shown correspond to bases 3 to 234 from the 3′ terminus of the primer. Reference values for this sequencing run are established in Fig. 5, which illustrates the principles discussed above for base assignment. Regions of the sequence are shown in greater detail in Fig. 7.

Improvements in sequencing. As illustrated by recent discussions of human genome sequencing, large-scale sequencing projects would require massive resources with current technology (*7*). Programs of intermediate scope, crucial in fundamental research as well as in the development of DNA probes, clinical reagents, and useful genetically engineered organisms, are likewise limited by the speed and ease of DNA sequencing.

Efforts to improve the technology of radioisotopic sequencing have been reviewed (*8*). In one example, a technique of continuous DNA blotting during electrophoresis is described in which the DNA fragments resulting from conventional dideoxy sequencing are separated by electrophoresis and allowed to elute from the bottom of the gel onto a moving membrane (*9*). The membrane is subsequently exposed to film and the resulting autoradiographic band patterns are interpreted in the normal fashion. This technique offers some advantages in readability of the band patterns and in the potential for a degree of automation in sequence analysis. However, the inherent disadvantages of using isotopically labeled reporters in the traditional manner remain and will limit the ultimate utility of this approach. A newer method with potential for genomic sequenc-

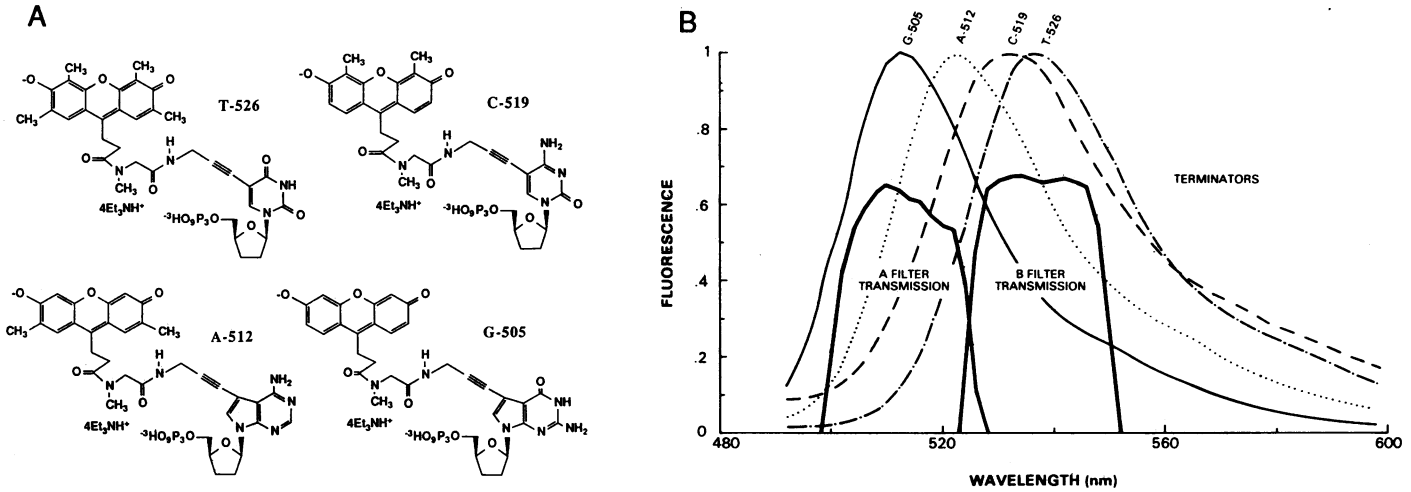

Fig. 2. Fluorescence-tagged chain-terminating reagents. (**A**) Chemical structures of the reagents used in modified dideoxy reactions for DNA sequencing. (**B**) Normalized fluorescence emission spectra of the reagents, measured in an 8 percent w/v polyacrylamide gel (19:1 w/w acrylamide:bis) under electrophoresis conditions similar to those of Fig. 6. The excitation wavelength was 488 nm. The absolute emission intensity values of the four compounds varied relative to each other by less than a factor of 2. Superimposed on the emission spectra are transmission functions of the interference filters used in the fluorescence detection system (Fig. 4). The nucleotide base assignment for each band is achieved by measuring the relative fluorescence signal in two detectors with spectral responses defined primarily by these filter functions.

ing employing radioactively labeled probes has also been described (*10*). This technique may address the problem of low throughput by enabling multiple sequences to be read from a single gel. The method does reduce the amount of gel preparation and electrophoresis required for a given sequencing task; however, sequence data are still obtained from autoradiographs which must be analyzed and transcribed in separate steps.

Advances in fluorescent sequencing have also been reported. The synthesis of oligonucleotides labeled at 5′ ends with fluorescent labels and useful for dideoxy sequencing (*11*), and several systems based on fluorescent primer technology (*12–14*), have been described. These systems successfully reduce or eliminate a number of the deficiencies of the current radioisotope technology. However, the disadvantages of primer labeling remain. In addition to the polymerase pausing problem noted previously, the use of labeled primers requires that oligonucleotides be custom synthesized and purified for each set of sequencing reactions. This restricts the use of sequencing strategies which employ multiple priming sites or different cloning vectors.

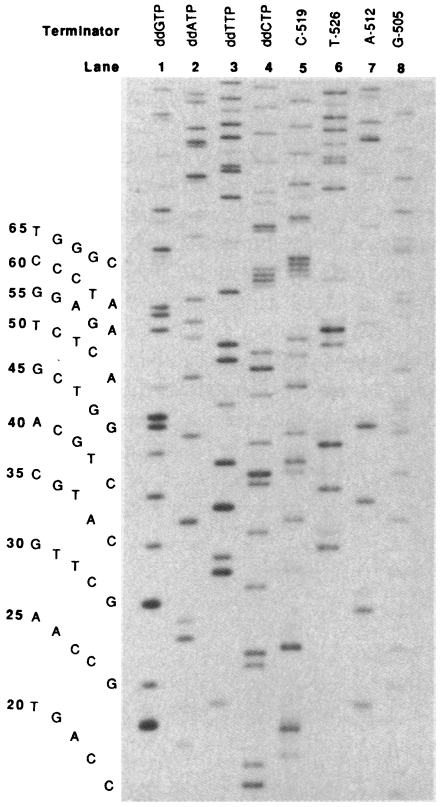

Fig. 3. Comparison of electrophoretic mobilities of DNA sequencing fragments terminated with ddNTP's and with fluorescence-tagged chain-terminators. The dideoxy method of Sanger (*2*) was used to sequence a region of M13*mp*18. A 17-bp oligonucleotide corresponding to the (−40) region of the template was end-labeled with [γ-^{32}P]ATP and polynucleotide kinase. Each reaction mixture contained template (0.1 μg), end-labeled primer (2.5 ng, annealed by heating to 95°C for 2 minutes and slowly cooled to room temperature), 60 m*M* tris-HCl at *p*H 8.5, 7.5 m*M* $MgCl_2$, 75 m*M* NaCl, 0.5 m*M* dithiothreitol, AMV reverse transcriptase (NEN, 20 units), and nucleotides at the following concentrations: lane 1, 2.5 μ*M* ddGTP, 12 μ*M* dGTP, 2.5 μ*M* dATP, 50 μ*M* dCTP and dTTP; lane 2, 0.25 μ*M* ddATP, 2.5 μ*M* dATP, 50 μ*M* dGTP, dCTP, and dTTP; lane 3, 3.0 μ*M* ddTTP, 12 μ*M* dTTP, 2.5 μ*M* dATP, 50 μ*M* dCTP and dGTP; lane 4, 2.5 μ*M* ddCTP, 12 μ*M* dCTP, 2.5 μ*M* dATP, 50 μ*M* dGTP and dTTP; lane 5, as in lane 4 except ddCTP has been replaced with 5.0 μ*M* C-519; lane 6, as in lane 3 except ddTTP has been replaced with 6.0 μ*M* T-526; lane 7, as in lane 2 except ddATP has been replaced with 2.0 μ*M* A-512; lane 8, as in lane 1 except ddGTP has been replaced with 8.0 μ*M* G-505. The primer extensions were carried out at 42°C for 10 minutes, and then unlabeled dNTP's (100 μ*M*) were added and the reaction proceeded for 10 minutes at 42°C. The reactions were stopped by diluting to 60 percent (v/v) formamide and denatured at 65°C for 10 minutes. A portion of each sample (1/8) was loaded onto an 8 percent (w/v) polyacrylamide gel containing urea (8*M*) and TBE buffer (100 m*M* tris-HCl, 83 m*M* boric acid, 1 m*M* Na_2EDTA, *p*H 8.1). The electrophoresis was carried out at 1600 volts.

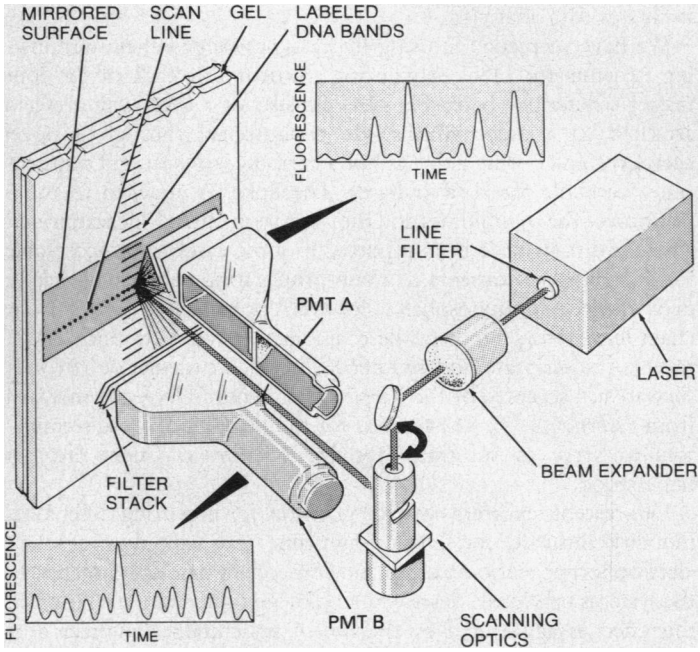

Fig. 4. Fluorescence detection system. Schematic drawing of the optical system used for scanning excitation and for measuring fluorescence from multiple sequencing lanes in an electrophoresis gel. Light from the argon ion laser is filtered to isolate the 488-nm emission line. The beam is deflected by a mirror into the scanning optics which are mounted on the shaft of a digitally controlled stepper motor. A lens focuses the beam into a spot in the plane of the gel. A second mirror directs the beam to a position on the scan line defined by the rotational position of the motor shaft. Sequencing of multiple samples is achieved by directing the beam sequentially to each of the sequencing lanes on the gel. Upon entering the gel, the beam excites fluorescence in the terminator-labeled DNA. Fluorescence is detected by two elongated, stationary photomultiplier tubes (PMT A and PMT B) which span the width of the gel. In front of each PMT, a filter stack is placed with one of the complementary transmission functions (Fig. 2B). Baseline-corrected ratios of signals in the PMT's are used to identify the labeled DNA fragments currently in the excitation region. Excitation efficiency and fluorescence collection are increased by the mirrored outer surface of the glass plate in the electrophoresis gel assembly.

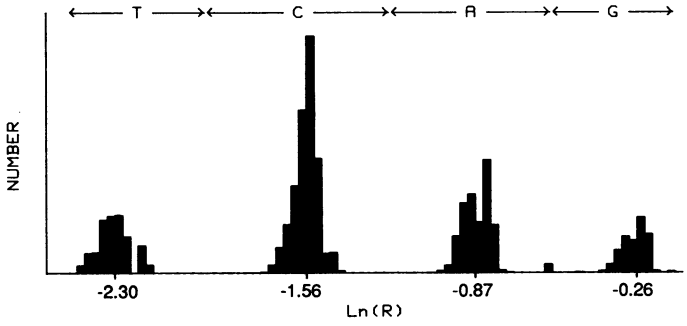

Fig. 5. Histogram of the 15 percent most certain ratios of baseline-corrected peak intensities of detector A to detector B for the run shown in Fig. 6. The natural logarithm of the ratio is used to enhance visual presentation. The ratios cluster about four reference values corresponding to the four bases as shown. The intervals used to assign the base identity of each peak are shown above the histogram. Data acquisition and analysis were performed by a Hewlett-Packard 9000 Model 500 computer.

The use of fluorescence-labeled chain-terminators for DNA sequencing, on the other hand, offers several distinct advantages. Labeled chain terminators afford complete flexibility of sequencing strategy and choice of vectors and allow the sequencing reactions to take place in one vessel. In this method, only true dideoxy terminations result in detectable fluorescent bands. Chain synthesis artifacts may still occur, but they will not be detected since these products are not fluorescent. This results in simplified elution patterns and facilitates data analysis.

We have succeeded in using fluorescence-tagged chain-terminating reagents for DNA sequencing. Structurally, each of the four reagents described here (Fig. 2A) consists of a succinylfluorescein attached to a dideoxynucleoside triphosphate through a novel acetylenic linker. The linker is both sterically compact and synthetically accessible for all four bases. The linker is attached to the 5 position of the pyrimidines and the 7 position of the 7-deazapurines. These positions have been reported in some cases to be acceptable for attaching substituents to chain-propagating substrates (that is, deoxynucleotide triphosphates) for DNA polymerases (*15, 16*). The chain-terminators reported here are incorporated by both AMV reverse transcriptase and modified T7 DNA polymerase (*6*), but they are not accepted by the Klenow fragment of DNA polymerase I from *Escherichia coli*. The reasons for the inability of these terminators to serve as substrates for the Klenow fragment are not understood.

Fluorescent reporters have the potential for imparting differential mobility shifts to the DNA fragments, thus complicating their electrophoretic elution pattern and subsequent base assignments. In the systems previously described employing more than one reporter, this effect is aggravated by the use of structurally dissimilar dyes (*12*). Perturbations to the linker arm structure and software algorithms were used to reconstruct the correct elution order of the fragments (*12*). Since the dyes reported here belong to a single family with only minor substituent differences on the succinylfluorescein moieties, they impart no significant differential mobility shifts to the DNA fragments and the correct elution order is observed in the raw data.

A fundamental criterion for an effective sequencing system is accuracy. Accuracy in detection and data analysis derives primarily from high sensitivity to the reporter, and in systems using multiple reporters, the ability to discriminate one label from another. A dideoxy sequencing system must be capable of detecting 10^{-15} to 10^{-16} mol of DNA per band (*12*). For a given number of reporter molecules, the measured fluorescence intensity is determined by the incident laser power, the optical properties of the reporter, and the efficiency of the detection system. This signal is superimposed on a background of Raman and elastic scattering arising from the interaction of the laser beam with the gel matrix and the glass plates of the electrophoresis assembly. Variation in this background determines the inherent noise in the total signal observed. Depending on lane geometry, gel thickness, and other factors, our system is capable of detecting, in real time, between 10^{-18} and 10^{-17} mol of succinylfluorescein reporter under sequencing conditions. Previously reported sensitivities for fluorescence-based sequencing systems are 3×10^{-18} mol (*14*) and from 10^{-17} to 5×10^{-17} mol (*12*).

Possible errors in sequencing include misidentifying a detected base, inserting a base, or failing to detect a base at all. To eliminate insertions and deletions, advantage can be taken of the uniform local temporal spacing of the DNA bands as they pass the detectors (*9*). The prediction of peak positions results in simplified detection and

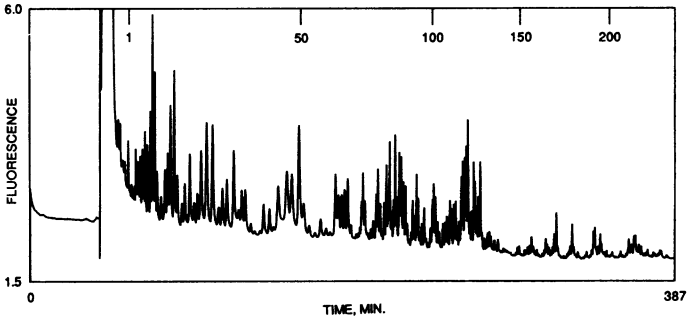

Fig. 6. Detection of fluorescent terminator-labeled DNA fragments from a region of M13*mp*18. Sum of the detector outputs for PMT A and PMT B (in arbitrary units of fluorescence intensity) versus time are shown for a 6-hour sequencing run. The fragments were generated in a two-stage reaction. In one reaction tube were added 3 μg of M13*mp*18 single-stranded DNA and 60 ng of primer (17 bp). This mixture was heated at 95°C for 2 minutes and then placed on ice for 5 minutes to anneal. In the primer extension reaction, 250 pmol each of dATP, dCTP, dTTP, and c^{7}dGTP were added and the reaction incubated at 42°C for 12 minutes in the presence of 60 m*M* tris-HCl, *p*H 8.3, 7.5 m*M* $MgCl_2$, 75 m*M* NaCl, 0.5 m*M* dithiothreitol, and 17 units of AMV reverse transcriptase. The extension reaction was stopped by the addition of a mixture containing 100 pmol of G-505, 800 pmol of A-512, 200 pmol of C-519, and 800 pmol of T-526 and then incubated for an additional 30 minutes at 42°C. Unincorporated fluorescent terminators were removed by gel filtration using a Sephadex G-25 spin column (5′ to 3′ Inc.). The effluent was dried under vacuum, washed with 70 percent ethanol, dried again, and resuspended in 5 μl of 90 percent formamide containing 11 m*M* Na_2EDTA. The sample was then heated to 65°C for 7 minutes and loaded onto an 8 percent w/v polyacrylamide gel (19:1 acrylamide:bis, 20 cm by 40 cm by 0.3 mm) containing 7*M* urea, 100 m*M* tris-HCl, *p*H 8.3, 83 m*M* boric acid, and 11 m*M* Na_2EDTA. The gel was electrophoresed for about 6 hours at 27 watts of constant power. Markers at top show the locations of bands corresponding to the bases.

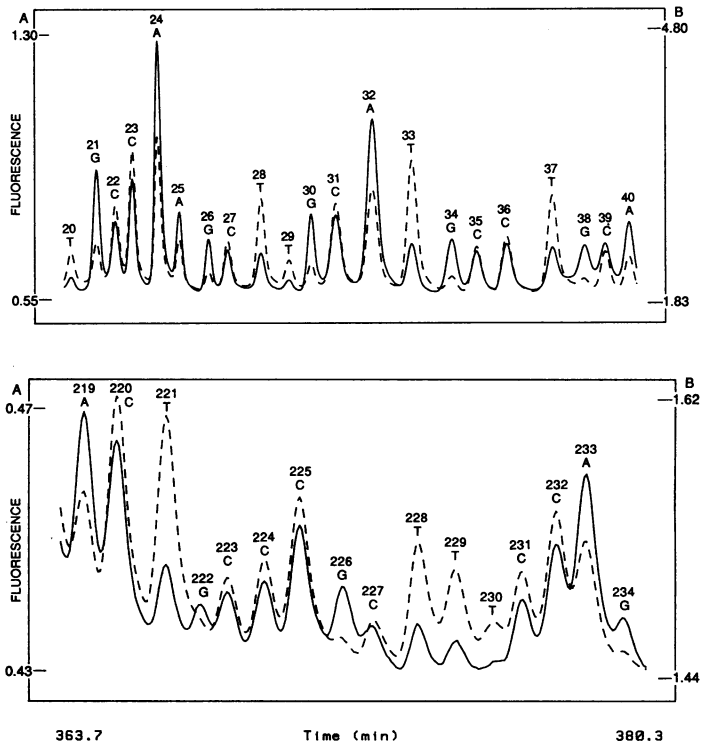

Fig. 7. Sequence of M13*mp*18 showing base assignments. (**Top**) Region of the sequence from base 20 to base 40. Each detector's output (in arbitrary fluorescence intensity units) is plotted as a function of time. Scale factors for the fluorescence are shown independently for the two detectors. Detector A, solid line, left scale; detector B, dashed line, right scale. (**Bottom**) Region of the sequence from base 219 to base 234. Similar data were obtained from base 3 to base 234.

identification of low-intensity peaks. The variation of detector signals across a band helps to identify peaks which are not fully resolved.

The nucleotide sequence of the M13*mp*18 region shown in Fig. 6 was determined by computer algorithms from base 3 to base 234 with no wrong base assignments and no missing or extra bases. Peak intensity variations result from sequence-dependent differences in rates of base incorporation and termination. This distribution is also affected by relative concentrations of dNTP's in the extension reaction and of the terminators in the final chain termination step, as well as by reaction times and conditions. More sequencing experience is required to determine the inherent accuracy in base assignment, but we are encouraged by the absence of errors for peaks above a minimum intensity threshold for discrimination.

In practice there are regions of DNA which are difficult to sequence due to aberrations in electrophoretic mobility caused by secondary structure (*17*). The data analysis system allows the location and extent of such regions to be identified so that flanking sequences remain in frame. 2′-Deoxy-7-deazaguanosine triphosphate has been used (c^7dGTP) in place of dGTP to minimize these effects. A similar technique has been used in fluorescent primer–based sequencing (*18*). For the run of Fig. 6, two regions of GC overlap were assigned by inspection.

A fundamental measure of the utility of an automated sequencer is raw throughput, defined as the sequencing rate per lane times the number of lanes per instrument. In our system 12 lanes are practical. After an initial period of electrophoresis during which the first DNA bands reach the detector, the sequencer is capable of determining approximately 50 bases per hour per lane. A fully loaded gel thus yields a throughput of about 600 bases per hour.

In summary, we have reported advances in chemistry and instrumentation for fluorescent labeling and detection of DNA fragments, and have described a rapid DNA sequencing sytem based on the use of fluorescent chain terminators. Further improvements in the performance of this system, as well as the emergence of other applications requiring sensitive DNA detection, are the subjects of current study.

REFERENCES AND NOTES

1. A. M. Maxam and W. Gilbert, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 560 (1977); *Methods Enzymol.* **65**, 499 (1980).
2. F. Sanger, S. Nicklen, A. R. Coulson, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463 (1977).
3. S. Biggs and F. G. Pope, *J. Chem. Soc. (London)* **123**, 2934 (1923).
4. For details on similar couplings see: M. J. Robins and P. J. Barr, *J. Org. Chem.* **48**, 1854 (1983).
5. J. L. Ruth and Y.-C. Cheng, *Mol. Pharmacol.* **20**, 415 (1981).
6. S. Tabor and C. C. Richardson, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 4767 (1987).
7. L. Roberts, *Science* **237**, 358 (1987); *ibid.*, p. 486.
8. W. J. Martin and R. W. Davies, *BioTechnology* **4**, 890 (1986).
9. S. Beck and F. M. Pohl, *EMBO J.* **3**, 2905 (1984).
10. G. M. Church and W. Gilbert, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 1991 (1984).
11. L. M. Smith, S. Fung, M. W. Hunkapiller, T. J. Hunkapiller, L. E. Hood, *Nucleic Acids Res.* **13**, 2399 (1985).
12. L. M. Smith, J. Z. Sanders, R. J. Kaiser, P. Hughes, C. Dodd, C. R. Connel, C. Heiner, S. B. H. Kent, L. E. Hood, *Nature (London)* **321**, 674 (1986).
13. W. Ansorge, B. S. Sproat, J. Stegemann, C. Schwager, *J. Biochem. Biophys. Methods* **13**, 315 (1986).
14. W. Ansorge, B. Sproat, J. Stegeman, C. Schwager, M. Zenke, *Nucleic Acids Res.* **15**, 4593 (1987).
15. R. M. K. Dale, D. C. Livingston, D. C. Ward, *Proc. Natl. Acad. Sci. U.S.A.* **70**, 2238 (1973).
16. P. R. Langer, A. A. Waldrop, and D. C. Ward, *ibid.* **78**, 6633 (1981).
17. N. L. Brown, *Methods Microbiol.* **17**, 259 (1984); N. S. Ambartsumyan and A. M. Mazo, *FEBS Lett.* **114**, 265 (1980); S. Mizusawa *et al.*, *Nucleic Acids Res.* **14**, 1319 (1986).
18. C. Connell *et al.*, *Biotechniques* **5**, 342 (1987).

19. We thank E. R. Adams, J. B. Clemens, D. E. Cope, G. C. Emmett, and R. K. Martin for technical assistance and our colleagues for support that made this research program possible. We especially acknowledge the encouragement provided by M. L. Pearson, F. F. Huppe, and C. T. Sciance.

4 September 1987; accepted 25 September 1987