FINAL REPORT
Ultrasensitive Fluorescence Detection of DNA Sequencing Gels
DE FG03 88ER60706
PI: Richard A. Mathies

DOE/ER/60706--1

DE92 001882

During the three years of this grant we have: (1) Developed and applied a new theory for optimizing high-sensitivity fluorescence detection. (2) Developed and patented a new high-sensitivity confocal-fluorescence laser-excited gel scanner. (3) Applied this scanner to the development of a new class of versatile and sensitive fluorescent dyes for DNA detection. (4) Developed methods for the detection of single fluorescent molecules by fluorescence burst detection. These results are summarized in more detail below.

A. Optimization of Laser-Excited Fluorescence

Our work on single molecule fluorescence detection of phycoerythrin (see below) led us to critically examine the optimization of laser excited fluorescence detection in the presence of ground state depletion and photodestruction. These single molecule detection experiments are feasible but difficult, so careful optimization is required. We needed a theory which would permit us to select the ideal illumination intensity and illumination time. Such a theory would not only facilitate single molecule detection but would help in the design and execution of all fluorescence detection experiments. The theory describe below has now been published in Analytical Chemistry (5) and several brief presentations of this work have been given (2, 4).

We consider a solution of fluorophores flowing through a focused laser beam. The emission from the illuminated volume consists of fluorescence from molecules passing through the beam superimposed on a continuous background due to Rayleigh and Raman scattering from the solvent and cell. To detect this fluorescence it is first necessary to make the fluorescence as intense as possible relative to the background. In the search for the optimal light intensity and transit or illumination time, we realized that the previous treatments of photodestruction assumed that the photodestruction rate and the rate of fluorescence are linearly related to the incident light intensity. This is true for weak illumination conditions but will not be true for the high light intensity conditions that occur as one tries to get the very best signal-to-noise ratio. At high light intensities a significant fraction of the molecules will be pumped to their excited states causing ground state depletion. This will alter the signal-to-noise ratio in the experiment by reducing the fluorescence-to-background ratio. The optical pumping should be intense but once the transition is saturated, increasing the light intensity further will just increase the background. Also, illuminating the molecule for a longer time will generate more fluorescent photons per molecule. However, there is no benefit in looking once the molecule has been photodestroyed. Our theory for selecting the optimal illumination conditions and transit time is outlined below.

The kinetic scheme that we will use to describe the fluorophores is given in Fig. 1. The fluorophore in the ground state S_0 is excited to S_1 by a photon, and the rate of absorption is given by

$$k_a = \sigma_a I = 3.8 \times 10^{-21} \epsilon I \quad (1)$$

where I is the light intensity (photon $\text{cm}^{-2} \text{s}^{-1}$), σ_a is the absorption cross section ($\text{cm}^2/\text{molecule}$) and ϵ is the molar extinction coefficient. S_1 can decay by emitting a photon k_f^0 , through nonradiative decay k_n , or through photochemistry k_d . The observed fluorescence decay rate k_f is determined by the sum of these processes.

$$k_f = k_f^0 + k_n + k_d = 1/\tau_f \quad (2)$$

Here τ_f is the observed fluorescence decay rate. Also, it is useful to define the intrinsic photodestruction time τ_d as $1/k_d$. The fluorescence quantum yield Q_f and the photodestruction quantum yield Q_d are given by

MASTER

$$Q_f = k^0_f/k_f \quad \{3\}$$

$$Q_d = k_d/k_f \quad \{4\}$$

We consider a solution flowing with velocity v (cm/s) through an exciting laser beam having a uniform intensity profile in a square of side L . The transit time for the molecule through the beam is $\tau_t = L/v$. Under these conditions, the number of photons emitted by the fluorophore during its transit through the beam is given by

$$n_f = (Q_f/Q_d) \{1 - \exp(-k\tau/(k+1))\} \quad \{5\}$$

where we have introduced the dimensionless variables $k = k_a/k_f$ and $\tau = \tau_t/\tau_d$. Meanwhile the scattering background is given by

$$n_b = (Q_f/Q_d) k \tau \alpha \quad \{6\}$$

where α is the ratio of the background scattering to fluorescence at very low light intensities and short illumination times. Note that n_b increases linearly with k and τ while n_f is intrinsically nonlinear.

Equations 5 and 6 provided the fundamental relations that permit us to optimize the conditions for fluorescence detection. Figure 2 presents a plot of $n_f/\sqrt{n_b}$ which represents the signal-to-background-noise ratio when background noise is dominant. The figure makes it clear that the S/N increases rapidly as k and τ are increased, but as k and τ approach unity an approximate plateau is reached. It is not practical to increase the transit time much beyond the photodestruction time because the S/N does not increase very rapidly. Similarly increasing the light intensity beyond that necessary to get the absorption rate equal to the fluorescence decay rate ($k=1$) is not generally valuable.

Although not summarized here, our paper also considers the possible effect of triplet saturation and/or photochemistry on the signal-to-noise ratio in these experiments (5).

B. Confocal Laser Gel Scanner for DNA Sequencing

The approach presented above is also useful for optimizing the fluorescence detection of DNA sequencing gels. A variety of methods for fluorescence detection of sequencing gels have been presented which have used either fluorescent labeled primers or terminators. Most commercial apparatus has employed a dedicated laser excitation system where the DNA bands are electrophoresed past the detection system. In our own work we have developed the off-line laser scanner system depicted in Figure 3. The gel is mounted on a motor driven scan stage. The argon ion laser beam is focused on the gel with a high numerical aperture microscope objective and then the fluorescence is gathered by the same objective and passed to the photodetector by a dichroic beam splitter. The advantages of this apparatus are: (1) The detection system is more effectively used because it is not dedicated to the electrophoresis system. A gel can be scanned in about one hour while the dedicated ABI sequencer requires ~10 hours for a run. (2) The optimum light intensity and transit times as defined above can be readily utilized. (3) The optical gathering of the fluorescence is very efficient in this retro-optical configuration. (4) The images have intrinsically very high ~10 micron resolution. (5) Most importantly, the use of a confocal excitation and detection geometry provides dramatic improvement in stray light rejection, in the rejection of scattered light from the glass substrates, and in the rejection of fluorescence from the substrate. Only the interior of the gel is effectively excited and the confocal detection ensures that only the fluorescence from the center of the gel is efficiently detected. This results in dramatic reductions of the background. Because of these fundamental advantages, we have filed and will receive a patent on the use of confocal detection in laser scanning of gels. This application has been licensed and is now under commercial development.

Figure 4 presents an image of a DNA sequencing gel detected with the laser scanner system. A fluorescein labeled primer was used to perform conventional Sanger sequencing on M13 DNA. The sequencing lanes were run at two different loads to illustrate the sensitivity limits of our apparatus. The lanes on the left were run at about 5 attomoles/band and those on the right were run at 1 attomole/band. The latter case corresponds to about 6×10^5 molecules per band. This is an excellent detection limit. This gel was run to determine the ultimate limit of the sensitivity of our system because this provides a good way to compare different detection systems. The sensitivity of other fluorescence detected sequencing systems are generally lower, ranging from 10^6 to 10^8 molecules per band.

Since the sensitivity and spatial resolution of our system are excellent, we decided to explore the miniaturization of the sequencing gels to increase the information density. Figure 6 shows gels that we have run using 2 mm lanes with an overall gel format of 12 cm x 15 cm. These images suggest that it should be feasible to reduce the lane size to ~1 mm. With a scan rate of 10 cm/s and 100 micron pixels, this image can be acquired in ~2 minutes. Enhanced resolution would be provided by doubling the electrophoresis length to 30 cm. A complete description of this apparatus and method for DNA sequencing has now been published in a detailed article in *BioTechniques* (6).

C. Confocal Detection of DNA Fingerprinting Gels and New Fluorescent Probe Development

Intercalation Complexes of DNA with Ethidium Homodimer. We have recently demonstrated that double-stranded DNA forms an unusually stable complex with ethidium homodimer cation (EthD) whose structure is shown in Figure 5 (3). The strong binding sites for EthD saturate at a ratio of one dye per four-to-five base pairs. Strong interaction between EthD and DNA had been reported more than ten years ago, but the extraordinary property of the complex discovered in our study is that it is stable in agarose gels under the usual conditions of electrophoresis. This stability makes possible the formation of the complex prior to electrophoresis followed by separation and detection in the absence of background stain. After electrophoretic separation, detection of bands containing 25 picograms of DNA was readily achieved in 1 mm thick agarose gels using laser excitation at 488 nm and the scanning confocal fluorescence imaging system described above. The band fluorescence intensity was linear with the amount of DNA applied to the gel from 0.2 to 1.0 nanogram per lane and with the number of kilobasepairs (kbp) per band within a lane. This procedure is at least two orders-of-magnitude more sensitive than conventional fluorescence-based procedures for DNA detection. We demonstrated that the method was effective in detecting and assessing the complexity of products formed by the polymerase chain reaction. These results have been described in detail in *PNAS* (3) and in *Nucleic Acids Research* (7).

Two-color DNA detection. We have also developed an optical system for 2-color confocal fluorescence scanning of DNA sequencing and mapping gels. The technical description of this apparatus will be published in *Methods in Enzymology* (8). The design is essentially the same as that used for our previously published one-color scanner (6) except that a dichroic beam splitter is used in the detection optical path to separate the two colors followed by two confocal detection systems. In the sequencing and mapping applications, both fluorophores are efficiently excited with the 488 nm argon ion laser line and have distinctive emission spectra. We have developed a novel two-color system of DNA detection on agarose gels by using ethidium homodimer and thiazole orange homodimer (Fig. 7). The size ladder is labeled with ethidium homodimer which forms a stable intercalation complex with the double-stranded DNA. Then, the unknown DNA fragments are labeled with the new dye, thiazole orange dimer (TOTO). Like ethidium homodimer, TOTO forms intercalation complexes with DNA that are stable to electrophoresis, but TOTO emits yellow fluorescence while EthD emits in the red. Thus, an unknown DNA sample can be labeled with TOTO and then run and detected in the same lane as the size standard. Two-color detection permits much more accurate determination of molecular weight because the standard and the

unknown are run in the same lane and detected at the same time. TOTO has the additional advantages that it is nonfluorescent in solution, its maximum absorption is at 488 nm, and its extinction coefficient is much greater than that of ethidium. This makes TOTO the most sensitive fluorescent DNA dye we have found, giving detection limits of a few picograms per band on agarose gels. These two-color sizing results are reported in our Methods in Enzymology article (8).

D. Single-molecule Fluorescence Detection

Figure 8 presents our experimental apparatus for single-molecule fluorescence detection (1). B-Phycoerythrin was selected as the fluorophore because of its high extinction coefficient at 514.5 nm, its high fluorescence quantum yield, and its large Stokes shift. To optimize the experiment based on the results in Figure 2, it is only necessary to know the fluorescence decay time τ_f and the intrinsic photodestruction time $\tau_d = 1/k_d$. First, the transit time is selected based on our determination that the intrinsic photodestruction time for B-PE is $\sim 700 \mu\text{s}$ (1). Reference to Figure 2 indicates that a τ of ~ 1 is needed to get up on the S/N plateau, so we selected a transit time of $800 \mu\text{s}$. Similarly, the excitation intensity was selected to give a value of k_a close to the average fluorescence lifetime of 2.5 ns ($k \sim 1$). Our theory tells us that these conditions will provide near optimum S/N for single molecule detection.

For single molecule detection the size of the imaged-illuminated volume must be optimized. The laser is tightly focused ($5\text{-}8 \mu\text{m}$ dia.) to lower the probability of multiple molecules in the probed volume, and a spatial filter is used to limit the imaged height to $2.5 \mu\text{m}$. This gives a very small 0.12 pL imaged volume. Such a small volume is necessary to ensure that we are in the single molecule occupancy limit. Even with this low volume it is necessary to run at extremely low concentrations ($<10^{-12} \text{ M}$) to guarantee that multiple occupancies of the probed volume will not occur. An advantage of tight focusing is that it fundamentally improves the fluorescence-to-background-noise ratio. With a very small beam, saturating light intensity can be achieved with very low overall powers. This reduces the background signal since the background scales linearly with the total incident power.

An equally important factor in single molecule detection is designing electronics to perform "single molecule time and amplitude gating". We have recently received a patent on this method and on a variety of electronic detector configurations for single molecule/single particle counting. The output from the photomultiplier is amplified and sent to a circuit that performs a running integration of the signal. This signal is then subjected to a series of analog comparator circuits that determine if the pulse has the height and width expected for a fluorescent molecule passing through the beam. With this type of "single molecule gating", the fluorescence signal/background ratio can be improved tremendously.

Figure 9 presents the distribution and autocorrelation functions for blank, PE monomers, and PE dimers. The blank closely follows a Poisson distribution with a mean value of 0.98 per 100 ms, and the autocorrelation shows random noise at all times. The distribution function for PE monomer shows marked deviation from a Poisson distribution. The fluorescent hump in the distribution plot for the dimer is more obvious, and it extends to higher counts rates. The evidence of single molecule detection is revealed by the autocorrelation analysis. The autocorrelation traces clearly show the presence of clustered bursts of emission having a duration of $\sim 700 \mu\text{s}$ characteristic of single molecule events. If we are indeed seeing single-molecule events, then the emission rate k_f should be twice as large for the dimer. This is in fact seen in Figure 9. Figure 10 presents the log of the number of single-molecule events versus the log of the concentration. The regression lines of these two plots have a slope fairly close to 1. The linear concentration dependence also demonstrates that we are seeing single-molecule events. Note that we can detect PE using our single molecule burst detector down to 10^{-15} M .

PUBLICATIONS

1. Peck, K., Stryer, L., Glazer, A. N., and Mathies, R.A. Single-Molecule Fluorescence Detection: Autocorrelation Criterion and Experimental Realization with Phycoerythrin, *Proc. Natl. Acad. Sci. U.S.A.* 86, 4087-4091 (1989).
2. Mathies, R. A., Peck, K., and Stryer, L. High-Sensitivity Single-Molecule Fluorescence Detection in Theory and Practice, *Proceedings of the 6th Annual BIOTECH USA Conference*, San Francisco CA, Oct. 2-4, 205-213 (1989).
3. A. N. Glazer, K. Peck and R. A. Mathies, A Stable Double-Stranded DNA-Ethidium Homodimer Complex. Application to Picogram Fluorescence Detection of DNA in Agarose Gels, *Proc. Natl. Acad. Sci. U.S.A.* 87, 3851-3855 (1990).
4. Mathies, R. A., Peck, K. and Stryer, L. High-sensitivity Single-Molecule Fluorescence Detection, *in* *Bioimaging and Two-dimensional Spectroscopy*, *Proceedings of the International Society for Optical Engineering-SPIE 1205*, 52-59 (1990).
5. Mathies, R. A., Peck, K. and Stryer, L. Optimization of High-Sensitivity Fluorescence Detection, *Anal. Chem.* 62, 1786-1791 (1990).
6. Quesada, M. A., Rye, H. S., Gingrich, J. C., Glazer, A. N., and Mathies, R. A. High-Sensitivity DNA Detection with a Laser-excited Confocal Fluorescence Gel Scanner, *BioTechniques* 10, 616-625 (1991).
7. Rye, H. S., Quesada, M. A., Peck, K., Mathies, R. A. and Glazer, A. N. High-Sensitivity Two-Color Detection of Double-Stranded DNA with a Confocal Fluorescence Gel Scanner Using Ethidium Homodimer and Thiazole Orange, *Nucleic Acids Research* 19, 327-333 (1991).
8. Rye, H. S., Yue, S., Quesada, M. A., Haughland, R. P., Mathies, R. A. and Glazer, A. N., Picogram Detection of Stable Dye-DNA Intercalation Complexes with a Two-Color Laser-Excited Fluorescence Gel Scanner, *Methods Enzymology*, in press.

PATENTS

1. U. S. Patent 4,979,824, December 25, 1990, High Sensitivity Fluorescent Single Particle and Single Molecule Detection Apparatus and Method K. Peck, R. A. Mathies and L. Stryer.
2. Laser Excited Confocal Fluorescence Gel Scanner, K. Peck and R. A. Mathies, patent filed January 12, 1990 (UCB 90-001) and now accepted.
3. Multichromophore Fluorescent Probes using DNA Intercalation Complexes, A. Glazer, K. Peck and R. A. Mathies, patent filed March 15, 1990 (UCB B90-002) and now accepted.

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Figure 1. Kinetic scheme for the ground S_0 , the first excited S_1 , and the triplet state T levels of a fluorophore. k_a is the rate of absorption, k_f^o is the natural radiative rate, and k_n is the radiationless decay rate.

Figure 2. Plot of $n_f/\sqrt{n_b}$ as a function of the relative light intensity k and relative transit time τ for a background intensity α of 1. The optimum value at each τ is indicated by the dots.

Figure 3. Diagram of a laser gel scanner for DNA sequencing. The stage has a 12" x 12" scan range with a resolution of ~ 10 micron. The typical laser beam size is 5-10 micron and the power is 2-5 mW at the sample. Typical scan rates are ~ 10 cm/s and the gel images shown below are digitized in ~ 100 micron pixels with ~ 100 micron between scanned lines.

Figure 4. Image of a fluorescence detected DNA sequencing gel. M13 DNA was sequenced using the Sanger method and fluorescein labeled primers. The left set of lanes were run at 5 attomole labeled fragments per band and the right set was run at 1 attomole per band. Our detection limit is about 6×10^5 molecules or one attomole per band.

Figure 5. Structures of ethidium homodimer and thiazole orange

Figure 6. Runs exploring the reduction of lane size and electrophoresis run length. The concentration is about 10^7 molecules per band and the lane size is 2mm. The total electrophoresis length was 12 cm and the gel was scanned at ~ 100 micron resolution.

Figure 7. Demonstration of two-color fluorescence detection on agarose gels with ethidium homodimer and thiazole orange. Sample: (1) linearized M13 stained with ethidium homodimer (2) linearized pBR322 stained with thiazole orange (3) M13 stained with ethidium + unstained pBR322 (4) M13 stained with ethidium homodimer + pBR322 stained with thiazole orange.

Figure 8. (A) Illustration of the concept of single molecule fluorescence detection. When a fluorescent molecule flows through the laser beam a burst of photons is generated; single-molecule events can be distinguished from the Poisson distributed background if the fluorescence emission rate is higher than the background emission rate. (B) Experimental apparatus for single molecule fluorescence detection. The 514.5 nm output from an argon ion laser is focused to a 8 μm diameter spot in a capillary tube through which a sample solution is flowed. The resulting fluorescence is collected at 90° and passed through spatial and spectral filters to the detector.

Figure 9. Comparison of burst height distribution (Left) and autocorrelation traces (Right) for blank (50 mM NaCl), PE monomer (2×10^{-13} M), and PE dimer (1×10^{-13} M). The solid lines are calculated Poisson distribution curves with a mean count rate of 0.98 per 100 ms. The number of data points in each analysis is 1.68×10^6 .

Figure 10. Fluorescence detection limits using single molecule counting. (A) PE monomer with concentrations from 2×10^{-16} to 2×10^{-12} M; (B) PE dimer with concentrations from 1×10^{-16} M to 1×10^{-12} M. To reduce the number of false events, the discriminator threshold was set so that only ~15% of the actual monomer events and 20% of the dimer events were detected.

END

**DATE
FILMED**

12/02/91