

Title: Progress Towards DNA Sequencing at the Single Molecule Level

RECEIVED
DEC 14 1995
OSTI

Author(s): Peter M. Goodwin, Rhett L. Affleck, W. Patrick Ambrose, James N. Demas, James H. Jett, John C. Martin, David S. Semin, Jay A. Schecker, Ming Wu, and Richard A. Keller

Submitted to: Proceedings of the International Workshop on Single Molecule Detection, Berlin, Germany, October 4-6, 1995; to appear in the journal "Experimental Technique of Physics"

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.



Los Alamos
NATIONAL LABORATORY

Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by the University of California for the U.S. Department of Energy under contract W-7405-ENG-36. By acceptance of this article, the publisher recognizes that the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. The Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy.

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED 

Form No. 836 R5
S 26-3 10/91
MASTER

Progress Towards DNA Sequencing at the Single Molecule Level

Peter M. Goodwin*[†] Rhett L. Affleck[†] W. Patrick Ambrose[†]
James N. Demas[‡] James H. Jett[§] John C. Martin[§]
Linda J. Reha-Krantz[¶] David J. Semin[†] Jay A. Schecker[†] Ming Wu[†]
Richard A. Keller[†]

Center for Human Genome Studies
Los Alamos National Laboratory, Los Alamos, New Mexico USA

October 30, 1995

ABSTRACT

We describe progress towards sequencing DNA at the single molecule level. Our technique involves incorporation of fluorescently tagged nucleotides into a targeted sequence, anchoring the labeled DNA strand in a flowing stream, sequential exonuclease digestion of the DNA strand, and efficient detection and identification of single tagged nucleotides. Experiments demonstrating strand specific exonuclease digestion of fluorescently labeled DNA anchored in flow as well as the detection of single cleaved fluorescently tagged nucleotides from a small number of anchored DNA fragments are described. We find that the turnover rate of *Escherichia coli* exonuclease III on fluorescently labeled DNA in flow at 36° C is ~7 nucleotides per DNA strand per second, which is approximately the same as that measured for this enzyme on native DNA under static, saturated (excess enzyme) conditions. Experiments demonstrating the efficient detection of single fluorescent molecules delivered electrokinetically to a ~3 pL probe volume are also described.

* Author to whom correspondence should be addressed

[†] Chemical Science and Technology Division.

[‡] Department of Chemistry, University of Virginia, Charlottesville, Virginia, USA

[§] Life Sciences Division.

[¶] Department of Biology, University of Alberta, Edmonton, Alberta, Canada.

1 Introduction

We describe recent experimental progress towards the goal of rapid, flow-based sequencing of long DNA fragments at the single molecule level. This technique will allow the rapid determination of long (~ 40 kilo-base pair) DNA sequences.¹⁻³ Briefly, the method involves: (i) the incorporation of fluorescently tagged nucleotides into a targeted sequence, (ii) anchoring of the tagged DNA strand in a flowing stream, (iii) exonuclease cleavage of the tagged nucleotides from the free end of the DNA strand, and (iv) detection and identification of single cleaved nucleotides by laser-induced fluorescence. The current status of our progress in each of these areas is summarized in Section 2 of this paper.

We describe the efficient detection of single tetramethylrhodamine isothiocyanate (TRITC) molecules. We used electrokinetic sample delivery from a pulled micropipette located in the sheath flow $\sim 50 \mu\text{m}$ upstream of the detection volume. These experiments are important steps towards the goal of flow-based DNA sequencing. Sample introduction from a micropipette simulates the delivery of cleaved nucleotides from the free end of an anchored DNA strand stretched out in flow. We have also detected single TRITC-labeled nucleotides cleaved from as few as 15 to 30 fluorescently tagged DNA strands anchored in flow. This is a major step towards proving the feasibility of our flow-based DNA sequencing method.

2 Current Status

2.1 Synthesis of Fluorescently Labeled DNA

A number of requirements are critical for the success of the method. First is the accurate incorporation of fluorescently tagged nucleotides into a targeted sequence. This is accomplished

by enzymatic synthesis of the complementary strand of the targeted DNA sequence using fluorescently labeled nucleotides as substrates. Initial work demonstrated the replication of single stranded M13 DNA out to 5000 to 7000 base pairs with two fluorescently tagged nucleotides (rhodamine-dCTP and fluorescein-dUTP) using T7 polymerase.³ The probability for misincorporation of the fluorescently labeled nucleotides is in the range of 10^{-5} to 10^{-4} per site.³ More recent work has demonstrated the complete replication of M13 DNA with two fluorescently tagged nucleotides as well as the replication of M13 DNA out to several thousand base pairs with three labeled nucleotides (fluorescein-dUTP, rhodamine-dCTP, and rhodamine-dATP) using a mutated T4 polymerase. In principle, only two distinct fluorescent labels are needed for sequencing; the sequence is built up using different combinations of the two tags and DNA bases in separate runs.⁴

2.2 Handling Single DNA Fragments

Our technique requires the ability to manipulate and anchor a single fragment of the fluorescently labeled DNA in a flowing stream. We have demonstrated binding of small numbers of biotinylated λ -DNA fragments to 2.8 μm diameter streptavidinated microspheres. The mean number of bound DNA fragments per microsphere can be controlled by varying the concentration of biotinylated λ -DNA incubated with the microspheres during the binding step. Binding conditions can be adjusted to give a mean number of attachments per microsphere of less than one. The number of attachments per microsphere roughly follow Poisson statistics and the peak of the distribution is proportional to the DNA concentration.² The biotin-avidin bond is robust. We have placed λ -DNA bound to the microspheres in flowing streams and observed the DNA strands⁵ stretch out to their full extension ($\sim 16 \mu\text{m}$) at linear flow velocities of $\gtrsim 0.1 \text{ mm/sec}$.³ Flow velocities of at least 10 mm/sec can be sustained without breakage of the biotin-avidin bond or the DNA. Extension of the end of the DNA strand out into the flow stream away from

the relatively stagnant region immediately downstream of the microsphere is necessary to avoid mis-ordering of the released nucleotides by diffusion.⁶

2.3 Exonuclease Digestion of Fluorescently Labeled DNA

Earlier work showed that a number of different exonucleases are able to digest fluorescein and rhodamine labeled DNA under static conditions (no flow).³ Preliminary measurements of the cleavage rate of *Escherichia coli* exonuclease III (Exo III), a 3' → 5' exonuclease, on rhodamine-labeled, (ds) M13 DNA in flow at room temperature indicated a rate of ~ 3 nucleotides per DNA strand per second.² This rate is approximately that measured using native DNA under static, enzyme saturated conditions.⁷

2.4 Detection and Identification of Single Fluorescently Labeled Nucleotides

Efficient detection and identification of the fluorescently tagged nucleotides after they are released from the DNA strand is crucial to the success of our approach. Single fluorophore detection by laser-induced fluorescence of chromophores commonly used as labels has been demonstrated by a number of groups during the past five years.⁸⁻¹⁶ Two chromophores, Rhodamine-6G and Sulforhodamine 101, have been detected concurrently and identified at the single molecule level using two excitation lasers and two spectrally distinct detection channels.¹⁷ More recent work has focused on effects that limit the detection efficiency of single molecules in flowing sample streams.¹⁸⁻²⁰ Diffusion of the analyte out of the sample stream and photodecomposition of the analyte in the probe laser prior to its detection are particularly important. Li and Davis have demonstrated detection of single Sulforhodamine 101 molecules with a reported efficiency of ~80%. A micropipette with a submicron opening was used to introduce a dilute sulforhodamine

solution into a sheath flow $\sim 15 \mu\text{m}$ upstream of the probe laser.²⁰

3 Experimental Methods

3.1 Single Molecule Detection Apparatus

The excitation source was an extended cavity Rhodamine 110 dye laser, synchronously pumped by a mode-locked argon ion laser operated at 514.5 nm. The dye laser was tuned to 554 nm, near the absorption maximum of TRITC and the TRITC labeled nucleotides and focused to a e^{-2} diameter of $14 \mu\text{m}$ in the center of a square bore flow channel ($250 \times 250 \mu\text{m}^2$). Fluorescence was collected at 90° using a $\times 40$, 0.85 NA microscope objective. The fluorescence collection side wall of the flow cell was fabricated to a coverslip thickness to accommodate the short working distance ($370 \mu\text{m}$) of the microscope objective. A slit ($600 \times 1000 \mu\text{m}^2$) located in the image plane of the collection objective with its long axis oriented parallel to the flow axis limited the detection volume to $\sim 3 \text{ pL}$. A bandpass filter (575 nm center wavelength, 26 nm FWHM) was located behind the slit. A long working distance $\times 32$, 0.55 NA microscope objective focused the light transmitted through the slit and filters onto the active area ($100 \times 100 \mu\text{m}^2$) of a single-photon avalanche diode.²¹ The time-correlated single-photon counting electronics used to discriminate the fluorescence from the prompt scattered light background are described in detail elsewhere.²²

3.2 Flow System and DNA Handling

A schematic of the flow system used for the exonuclease digestion experiments is shown in Figure 1. Sheath buffers were delivered to the flow cell at a constant volumetric flow rate using a syringe pump. A four-way valve was used to select which buffer passed through the flow cell.

Fluorescent impurities present in the buffers were photolyzed prior to their introduction into the flow cell using an in-line photolysis cell.²³

A micropipette, pulled from a quartz capillary (240 μm O.D., 100 μm I.D.), was used to hold a single streptavidin coated microsphere to which DNA was bound in the flowing sheath stream. A microsphere was loaded by pushing the tapered end of the micropipette out of the square bore flow channel and inserting it into a droplet of buffer containing a suspension of microspheres. Suction applied to the back of the quartz capillary was used to capture a single microsphere. The micropipette and trapped microsphere were then pulled back into the flow channel. We were able to verify that a single microsphere was trapped by viewing the end of the micropipette through the $\times 40$ microscope objective used to collect fluorescence. Two views of a micropipette, with and without a trapped microsphere, are shown in Figure 3. The trapped microsphere was positioned approximately 50 μm upstream of the focused excitation laser beam. For measurements at elevated temperatures, the quartz flow cell was warmed using a thin film heater (not shown in Figure 1) attached to the flow chamber face opposite the collection side. After each experiment, a small thermocouple was inserted into the flow channel to measure the temperature of the fluid inside the flow cell.

3.3 Electrokinetic Sample Delivery

For some of the experiments, a smaller micropipette was used for electrokinetic sample delivery. This was a quartz capillary (90 μm O.D., 20 μm I.D.), approximately 15 cm long. One end of the capillary was pulled to a tip with a opening < 1 μm in diameter. The capillary was inserted into a larger capillary (240 μm O.D., 100 μm I.D.) in order to keep it approximately centered in the 250×250 μm^2 square bore flow channel. A solution of fluorescent dye (1 μM TRITC) dissolved in a salt buffer was delivered electrokinetically from the capillary tip into the

sheath flow. The back end of the capillary was connected to the positive electrode of a DC power supply through a salt bridge consisting of a 100 μm I.D., 15 cm long capillary filled with buffer. The power supply return was connected to the sheath buffer at the flow cell drain. The potential drop across the sample delivery capillary pumped the fluorescent dye with electrokinetic flow, the sum of the electroosmotic fluid velocity and the electrophoretic drift velocity of the fluorescent dye.²⁴ Figure 3 shows the capillary tip and sample stream fluorescence excited by the laser as viewed through the collection objective. Sheath fluid was delivered by the same flow system used for the exonuclease cleavage experiments (Figure 1), however, only one of the delivery paths was used. Again, fluorescent impurities present in the sheath buffer were photolyzed upstream of the flow cell.²³

3.4 Materials

Two sheath buffers were used for the exonuclease digestion experiments. The "inactive" buffer was 50 μM TRIS-HCl and 5 mM MgCl_2 dissolved in high purity water. The "active" buffer was identical to the inactive buffer except for the addition of Exo III exonuclease (New England BioLabs, Beverly MA) at a concentration of 300 units per mL. Two different sources of fluorescently labeled DNA were used in these experiments. The first sample (provided by Dr. J. Harding, Life Technologies Inc., Gaithersburg, MD) was from M13mp(+) single stranded vector that was replicated using three normal nucleotides (dATP, dGTP, and dTTP) and one fluorescently labeled nucleotide, TRITC-dCTP. The M13 vector was hybridized with a biotinylated primer. The complimentary strand was polymerized in a primer extension reaction using a modified T5 polymerase lacking 5' exonuclease activity resulting in fluorescently labeled double stranded DNA. This sample was not completely replicated, the size distribution (measured using gel electrophoresis) of the labeled DNA extended from a few hundred bases pairs (bp) up to ~ 7000 base pairs with a mean replication length centered between 2000 and 3000 base pairs.

Incubation of the sample with 10.5 μm diameter avidin coated latex microspheres was carried out with a large excess ratio of DNA fragments to microspheres to enhance the number of fragments bound to each microsphere. Approximately 50 to 100 fragments of DNA were bound to each microsphere.²

Linda, please describe your fluorescently tagged M13 DNA sample and the characterization of the distribution of tagged lengths here.

A second sample was biotinylated (ds) M13 DNA with the (-) strand completely (?) replicated using three normal nucleotides (dATP, dGTP, and dCTP) and fluorescent TRITC-dUTP.

This sample was incubated with 2.8 μm diameter streptavidin coated magnetic microspheres (Dynal Inc., Lake Success, NY). Approximately 15 to 30 DNA fragments were attached to each microsphere under the conditions used. This was determined by staining a sample of the microspheres with YOYO-1 (benzoxazolium-4-quinolinium dimer, Molecular Probes Inc., Eugene OR), a DNA intercalating dye, and viewing the microspheres under an epi-fluorescence microscope. The average number of bound DNA fragments per microsphere was estimated by counting the number of fragments observed on ~ 50 microspheres. Two views of YOYO-1 stained M13 DNA fragments bound to a microsphere are shown in Figure 4. For the electrokinetic sample delivery experiments, the sample was a 1 μM solution of TRITC (Molecular Probes, Inc.) dissolved in Dulbecco's Phosphate-Buffered Saline (1 \times D-PBS, GIBCO BRL). The salt bridge used a 1 \times D-PBS buffer. The same buffer diluted by a factor of 50 to 100 \times in high purity water served as the sheath fluid for the electrokinetic sample delivery.

4 Results

4.1 Efficient Detection of Single TRITC Molecules

We used electrokinetic delivery of TRITC to the probe volume to optimize the experimental parameters as well as to demonstrate efficient detection of this species at the single molecule level. Figure 4 shows the sample stream illuminated by the focused excitation beam under the flow conditions (sheath volumetric flow rate, $20 \mu\text{L}/\text{min}$; linear flow velocity, $\sim 1 \text{ cm}/\text{sec}$) used for the experiments. The elongated shape of the illuminated sample stream reveals that its diameter ($\lesssim 10 \mu\text{m}$) is significantly smaller than that of the focused excitation laser beam (e^{-2} diameter: $14 \mu\text{m}$). At electrokinetic drive voltages (voltage drop across the whole circuit, including the salt bridge) on the order of 1000 volts, the delivery rate of the TRITC was high enough to make the illuminated sample stream visible to the naked eye (as shown in the figure). Reduction of the voltage allowed sample delivery to the detection volume at extremely low rates. This is demonstrated in Figure 5 where we plot the number of delayed fluorescence photons detected in successive 0.25 millisecond time intervals for four different drive voltages. At 2.25 volts the delivery rate of TRITC molecules is still too high for single molecule detection—more than one molecule is in the probe volume at any given time. As the drive voltage is lowered to 1.5 volts, the delivery rate drops to the point where individual TRITC molecules can be seen transiting the probe volume. The similarity in the sizes of the photon bursts at the lowest drive voltage is the result of all TRITC molecules in the sample stream receiving approximately the same integrated excitation.

A more quantitative measure of the size distribution of fluorescence bursts due to TRITC delivered from such a capillary is shown in Figure 6. For this experiment, the average excitation laser power was 20 milliwatts, the sheath ($0.02\times$ D-PBS buffer) volumetric flow rate was 20

$\mu\text{L}/\text{min}$, and the electrokinetic drive voltage was ~ 2 volts. Approximately 64 seconds of data were sifted for photon bursts using a search algorithm described elsewhere^{11,22}; these results are shown as open circles in panel a) of the figure. The filled circles are the result of sifting approximately 64 seconds of blank data (taken with the capillary tip positioned to cause the sample stream to miss the probe volume); this distribution reflects the presence of fluorescent impurities in the sheath fluid that either survived the photolysis step or were introduced downstream of the in-line photolysis cell. The peak at ~ 40 photoelectrons (PE) is due to the passage of single TRITC molecules through the probe volume; the shoulder on the high side of the main peak extending past 100 PE is due to fluorescence bursts detected from two or more TRITC molecules passing through the probe volume simultaneously. The experiment was numerically simulated; the distribution shown in panel b) is the result of sifting 64 seconds of synthetic data for fluorescence bursts. The Monte Carlo simulation used to generate this data accounts for spatial variations in the intensity of the focused excitation laser, the light gathering efficiency of the optical collection system,²⁵ optical saturation, photobleaching and diffusion of the analyte molecules. The simulation used a mean delivery rate of 100 TRITC molecules per second from a source $1 \mu\text{m}$ in diameter located $50 \mu\text{m}$ upstream of the probe volume. Additional parameters used in the simulation are listed in Table 1. The good agreement between the experiment and simulation corroborates our assertion that we are detecting single TRITC molecules.

If a threshold for detection is set at 20 PE, the simulation predicts that $>90\%$ of the TRITC molecules leaving the capillary are detected. According to the simulation, approximately 10% of the TRITC molecules photobleach prior to crossing the probe volume; about half of these are detected. The measured background rate for bursts ≥ 20 PE was $\sim 3 \text{ sec}^{-1}$. The estimated detection rate of TRITC molecules was $\sim 100 \text{ sec}^{-1}$ (accounting for bursts from two molecules simultaneously present in the probe volume).

Table 1. Simulation Parameters

TRITC absorption cross-section at 554 nm ^a	$3.1 \times 10^{-16} \text{ cm}^2$
TRITC fluorescence quantum yield ²⁷	0.28
TRITC photodestruction quantum yield ²⁶	6×10^{-6}
TRITC diffusion coefficient ^b	$3.3 \times 10^{-6} \text{ cm}^2 \text{ sec}^{-1}$
TRITC saturation intensity ^c	$4 \times 10^4 \text{ Watt cm}^{-2}$
Average excitation laser power	$20 \times 10^{-3} \text{ Watt}$
Excitation laser beam e^{-2} diameter	14 μm
Sheath fluid linear flow velocity ^d	0.79 cm sec^{-1}
Absolute fluorescence photon detection efficiency	1×10^{-2}
Background count rate per Watt excitation power	$1.4 \times 10^5 \text{ sec}^{-1} \text{ Watt}^{-1}$

^aEstimated from the TRITC absorption spectrum and the reported²⁶ absorption cross-section at 515 nm.

^bValue measured for the dyes Eosin and Rhodamine-B dissolved in water.²⁸

^cTime averaged saturation intensity (intensity at which the fluorescence emission rate is one half of its saturated value) measured using the mode-locked dye laser.

^dDerived from the measured transit time (1.8 msec) and the excitation laser beam diameter.

4.2 Exonuclease Digestion of Fluorescently Labeled DNA in Flow

Preliminary experiments were carried out with fluorescently labeled M13 DNA fragments attached to 10.5 μm diameter streptavidin coated latex microspheres.² Buffers were delivered at a volumetric flow rate of 25 $\mu\text{L}/\text{min}$; the average sheath fluid velocity down stream of the microsphere was $\sim 0.7 \text{ cm}/\text{sec}$. The average excitation laser power was 10 milliwatts. The in-line photolysis step shown in Figure 1 was not used for these experiments. We estimate that 50 to 100 fluorescently labeled M13 fragments were bound to the microsphere. The experiment was done at room temperature. Data from one of these measurements is shown in Figure 7. The fluorescence (photons detected in a time window excluding the prompt scattered background) count rate is plotted against time. Initially, inactive buffer was flowed over the DNA anchored $\sim 50 \mu\text{m}$ upstream of the probe volume. At 60 seconds the four way valve was switched to admit active buffer to the flow cell. After a delay of ~ 300 seconds (due to the dead volume between the valve and the flow cell) the count rate started to increase as the exonuclease began to digest the anchored DNA. The gated photon count rate reached a maximum of 4000 sec^{-1} and then

decreased as the fluorescently tagged DNA fragments were digested. At 1560 seconds the valve was switched to admit inactive buffer to the flow cell. The count rate dropped again as the active buffer was swept out of the flow cell. This difference was due to the increased fluorescent impurities present in the active buffer compared to the inactive buffer. At 2400 seconds the valve was again switched to admit active buffer. Note the smaller rise in the fluorescence count rate – the fluorescently labeled DNA was completely digested. Finally, at 3400 seconds the valve was switched back to inactive buffer. This experiment demonstrates exonuclease cleavage of fluorescently tagged DNA in flow. The exonuclease cleavage rate under these conditions was estimated to be approximately 3 bases per DNA strand per second.² This rate is somewhat higher than that (1.7 bases per DNA strand per second) measured⁷ for Exo III on native DNA under static, enzyme saturated conditions. This discrepancy may be due to the uncertainty in the mean size of the fluorescent DNA fragments used in the experiment.

A number of changes were made to improve the exonuclease cleavage experiment. First, was the addition of the in-line photolysis step, shown in Figure 1 and described in the Experimental Methods section, to reduce the background due to fluorescent impurities present in the inactive and active buffers. Photolysis of the buffers reduced the background burst rate by approximately a factor of ten without reducing the activity of the Exo III exonuclease significantly.²³ Smaller 2.8 μm diameter microspheres were used to anchor the DNA in flow to decrease the diameter of the stream of cleaved nucleotides released into the sheath flow. Narrow sample streams allow more efficient detection of cleaved fluorescently labeled nucleotides. Fluorescently labeled M13 DNA fragments with both a larger average length and a narrower distribution of lengths were used. Finally, the flow cell was heated to 36° C to increase the exonuclease digestion rate. The buffer flow rate and excitation laser power were the same as those in the previous experiment. Results of one such experiment are shown in Figure 8. Panel a) is a plot of the fluorescence count rate versus time. Initially, inactive buffer was flowing through the cell; the four way valve was

switched to admit active buffer to the flow cell at time “A”. After ~ 1000 seconds the fluorescence count rate increased as the active buffer reached the DNA bound to the support. At time “B” the valve was switched to admit inactive buffer to the flow cell. The data was searched for photon bursts ≥ 5 PE in size. A time sequence of the detected burst rate is plotted in panel b). Each point is the burst rate derived from 30 seconds of data. The solid line is an estimate of the background burst rate due to the inactive and active buffers. Panel c) of the figure shows the time sequence of the burst rate after subtraction of the background.

Integration of the time sequence shown in panel c) of Figure 8 results in ~ 22900 detected bursts. Assuming these are due to individual cleaved fluorescently labeled nucleotides, a lower limit estimate of the number of M13 fragments bound to the support and accessible by the exonuclease is, $\frac{22900 \times 4}{6600} \simeq 14$. Here it is assumed that approximately one quarter of the cleaved nucleotides are fluorescently tagged (TRITC-dUTP), the TRITC-dUTP detection efficiency is unity, and each M13 fragment is labeled out to its full length, i.e., 6600 base pairs. This number is on the low end of the estimated number of fragments (15–30) seen on the microspheres using epi-fluorescence microscopy, however, some of these are not accessible by the exonuclease due to the interference of the micropipette holding the microsphere (Figure 3). Another possibility is that the sample stream diameter of cleaved nucleotides is still too large and/or not well aligned with the probe volume and hence the detection efficiency for fluorescently labeled nucleotides is significantly less than unity. We know that the sample stream of nucleotides released from the microsphere was either not as small and/or not as well aligned with the probe volume as that obtained using electrokinetic sample delivery from the pulled capillary. This is based on the observation that the burst size distribution from the cleaved fluorescently labeled nucleotides is not peaked away from zero PE as is the case for the pulled capillary (Figure 6). We expect, however, that the detection efficiency for cleaved fluorescently labeled nucleotides was > 0.5 based on the number of M13 fragments (15–30) seen on the microspheres.

An estimate of the exonuclease cleavage rate based on the peak burst rate ($\sim 25 \text{ sec}^{-1}$) and the number of DNA fragments (~ 14) from which cleaved fluorescent nucleotides were detected is, $\frac{25 \times 4}{14} \simeq 7$ nucleotides per DNA fragment per second. Again, it is assumed that approximately one quarter of the nucleotides are fluorescently labeled dUTP's. Note that this estimation of the cleavage rate does not depend of the single molecule detection efficiency – this factor drops out of the ratio used to calculate the rate. Our measured rate is approximately the same as that measured for this exonuclease on native DNA under static conditions at 36° C .⁷ Due to the narrower distribution of DNA fragment lengths used, we have more confidence in the exonuclease rate derived from this data compared to that obtained from the previous experiment.² Work is underway to measure the temperature dependence of the Exo III cleavage rate using the improved setup.

5 Summary

Each of the steps necessary for our flow based DNA sequencing method has been shown to work. In this paper we described experiments demonstrating the exonuclease cleavage of fluorescently labeled DNA in flow as well as the downstream detection of the fluorescently tagged nucleotides at the single molecule level. We have measured cleavage rates for *Escherichia coli* exonuclease III on fluorescently labeled DNA in flow in the range of 3 to 7 bases per DNA strand per second; approximately the same as reported⁷ for this exonuclease on native DNA under static conditions. We have also demonstrated the efficient detection of single TRITC molecules under conditions similar to those anticipated for flow based DNA sequencing.

We have detected cleaved fluorescently labeled nucleotide from as few as 14 fragments of labeled DNA. In order to reach our goal of detection of labeled nucleotides from a *single* strand of DNA, we need to increase the exonuclease cleavage rate by at least an order of magnitude or

reduce the background burst rate due to fluorescent impurities present in the buffers. In the near term, we will explore the possibility of using Exo III at relatively high temperatures ($\sim 60^\circ \text{C}$). The cleavage rate of this exonuclease under saturated (excess enzyme) conditions is expected to double for each 6°C temperature rise.⁷ Work is underway also to reduce the background burst rate due to fluorescent impurities present in the buffers and flow system.

6 Acknowledgments

We would like to thank Harvey Nutter for his expert technical assistance in the laboratory. This work was supported by the Los Alamos Center for Human Genome Studies under United States Department of Energy contract W-7405-ENG-36.

7 Figure Captions

Figure 1. Flow system for exonuclease digestion experiments. Buffers were delivered to the flow cell at a volumetric flow rate of $25 \mu\text{L}/\text{min}$. The mean fluid velocity inside the flow channel was $\sim 0.7 \text{ cm}/\text{sec}$.

Figure 2. A micropipette holding a $2.8 \mu\text{m}$ diameter microsphere is shown in panel a); panel b), the micropipette after the microsphere has been released.

Figure 3. Electrokinetic delivery of TRITC via a drawn micropipette. View of the capillary tip and sample stream fluorescence seen through the $\times 40$ collection objective. The sheath flow direction is from bottom to top in the picture. The end of the capillary is located $\sim 40 \mu\text{m}$ upstream of the excitation laser. Conditions: sheath volumetric flow rate, $20 \mu\text{L}/\text{min}$; sheath linear flow velocity, $\sim 1 \text{ cm}/\text{sec}$; electrokinetic drive voltage, $+1500 \text{ VDC}$.

Figure 4. Two views of YOYO-1 stained M13 (ds) DNA fragments bound to a $2.8 \mu\text{m}$ diameter streptavidin coated microsphere. The plane of focus is near the top of the microsphere in panel a) and near the center of the microsphere in panel b).

Figure 5. Detection of fluorescence from TRITC delivered electrokinetically from a drawn capillary, geometry similar to that shown in Figure 4. For this experiment, the average excitation laser power was 10 milliwatts, and the sheath ($0.01\times$ D-PBS buffer) volumetric flow rate was $25 \mu\text{L}/\text{min}$. Fluorescence counts (gated counts) detected in 0.25 millisecond bins over 200 millisecond intervals are shown for four electrokinetic drive voltages: a) 2.25 V, b) 2.00 V, c) 1.75 V, d) 1.50 V. Photon bursts due to individual TRITC molecules are clearly visible at the lowest drive voltage, panel d).

Figure 6. Experimental and simulated photon burst size distributions. Panel a), open circles, measured burst size distribution for TRITC delivered electrokinetically from a drawn capillary. Filled circles, measured background burst size distribution. Panel b), photon burst size distribution obtained from synthetic data, simulation parameters listed in Table 1. Each distribution was derived from 64 seconds of data.

Figure 7. Time history of the gated fluorescence count rate from a flowing exonuclease cleavage experiment; each data point is the rate calculated from 4 seconds of data. The legend below the fluorescence data indicates the times when the four way valve was switched between the inactive and active buffers. Approximately 50 to 100 fluorescently labeled M13 DNA fragments were bound to a 10 μm diameter streptavidin coated microsphere anchored ~ 50 μm upstream of the probe volume.

Figure 8. Detection of single fluorescently labeled nucleotides enzymatically cleaved from 15–30 fluorescently labeled M13 DNA fragments bound to a 2.8 μm diameter streptavidin coated microsphere anchored in flow. Panel a), time history of the gated fluorescence count rate; each data point is calculated from 4 seconds of data. The four way valve was switched to admit active buffer into the flow system at time “A”, inactive buffer was readmitted at time “B”. Panel b), detection rate of photon bursts ≥ 5 photoelectrons in size. The solid curve is the estimated background due to inactive and active buffers. Panel c), photon burst rate after subtraction of the background bursts. Each of the open circles in panels b) and c) is the burst rate calculated from 30 seconds of data.

8 REFERENCES

- [1] Jett, J. H., Keller, R. A., and ? *United States Patent # 4,962,037* (1990).
- [2] Schecker, J. A., Goodwin, P. M., Affleck, R. L., Wu, M., Martin, J. C., Jett, J. H., Keller, R. A., and Harding, J. D. *SPIE* **2386**, 4–12 (1995).
- [3] Ambrose, W. P., Goodwin, P. M., Jett, J. H., Johnson, M. E., Martin, J. C., Marrone, B. L., Schecker, J. A., Wilkerson, C. W., and Keller, R. A. *Ber. Bunsenges. Phys. Chem.* **97**(12), 1535–1542 (1993).
- [4] Jett, J. H., Keller, R. A., Martin, J. C., Posner, R. G., Marrone, B. L., Hammond, M. L., and Simpson, D. J. *United States Patent # 5,405,747* (1995).
- [5] . The DNA strand was stained with the intercalating dye benzothiazolium-4-quinolinium dimer (TOTO-1) obtained from Molecular Probes Inc., Eugene, OR 97402. The stained DNA was viewed under an epifluorescence microscope.
- [6] Pratt, L. R. and Keller, R. A. *J. Phys. Chem.* **97**, 10254 (1993).
- [7] Hoheisel, J. D. *Anal. Biochem.* **209**, 238–246 (1992).
- [8] Shera, E. B., Seitzinger, N. K., Davis, L. M., Keller, R. A., and Soper, S. A. *Chem. Phys. Lett.* **174**(6), 553–557 (1990).
- [9] Soper, S. A., Shera, E. B., Martin, J. C., Jett, J. H., Hahn, J. H., Nutter, H. L., and Keller, R. A. *Anal. Chem.* **63**, 432–437 (1991).
- [10] Soper, S. A., Mattingly, Q. L., and Vegunta, P. *Anal. Chem.* **65**, 740–747 (1993).
- [11] Wilkerson, C. W., Goodwin, P. M., Ambrose, W. P., Martin, J. C., and Keller, R. A. *Appl. Phys. Lett.* **62**(17), 2030–2032 (1993).
- [12] Mets, U. and Rigler, R. *J. Fluorescence* **4**(3), 259–264 (1994).

- [13] Nie, S., Chiu, D. T., and Zare, R. N. *Science* **266**, 1018–1021 (1994).
- [14] Nie, S., Chiu, D. T., and Zare, R. N. *Anal. Chem.* **67**(17), 2849–2857 (1999).
- [15] Barnes, M. D., Ng, K. C., Whitten, W. B., and Ramsey, J. M. *Anal. Chem.* **65**(17), 2360–2365 (1993).
- [16] Soper, S. A., Davis, L. M., Fairfield, F. R., Hammond, M. L., Harger, C. A., Jett, J. H., Keller, R. A., Marrone, B. L., Martin, J. C., Nutter, H. L., Shera, E. B., and Simpson, D. J. *SPIE* **1435**, 168–178 (1991).
- [17] Soper, S. A., Davis, L. M., and Shera, E. B. *J. Opt. Soc. Am. B* **9**, 1761 (1992).
- [18] Davis, L. M. and Li, L. Q. *OSA Technical Digest Series: Laser Applications of Chemical Analysis* **5**, 206–209 (1994).
- [19] Lee, Y. H., Maus, R. G., Smith, B. W., and Winefordner, J. D. *Anal. Chem.* **66**(23), 4142–4149 (1994).
- [20] Li, L. Q. and Davis, L. M. *Appl. Optics* **34**(18), 3208–3217 (1995).
- [21] Li, L. Q. and Davis, L. M. *Rev. Sci. Instrum.* **64**(6), 1524–1529 (1993).
- [22] Goodwin, P. M., Wilkerson, C. W., Ambrose, W. P., and Keller, R. A. *SPIE* **1895**, 79–89 (1994).
- [23] Affleck, R. L., Demas, J. N., Goodwin, P. M., Schecker, J. A., Wu, M., and Keller, R. A. . Manuscript in preparation.
- [24] Oda, R. P. and Landers, J. P. In *Handbook of Capillary Electrophoresis*, Landers, J. P., editor, chapter 2. CRC Press, Florida (1994).
- [25] Goodwin, P. M., Ambrose, W. P., Martin, J. C., and Keller, R. A. *Cytometry* **21**, 133–144 (1995).
- [26] Soper, S. A., Nutter, H. L., Keller, R. A., Davis, L. M., and Shera, E. B. *Photochemistry and Photobiology* **57**(6), 972–977 (1993).

- [27] Tsien, R. Y. and Waggoner, A. In *Handbook of Biological Confocal Microscopy*, Pawley, J. B., editor, chapter 16. Plenum Press, New York 2nd edition (1990).
- [28] Marenesco, N. S. *J. Chim. Phys.* 24, 593 (1927).

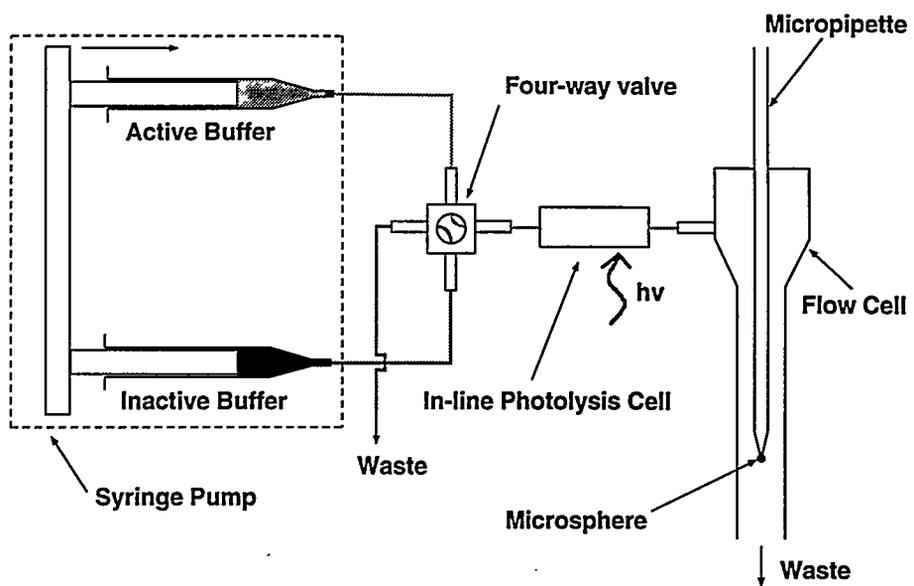


Figure 1:

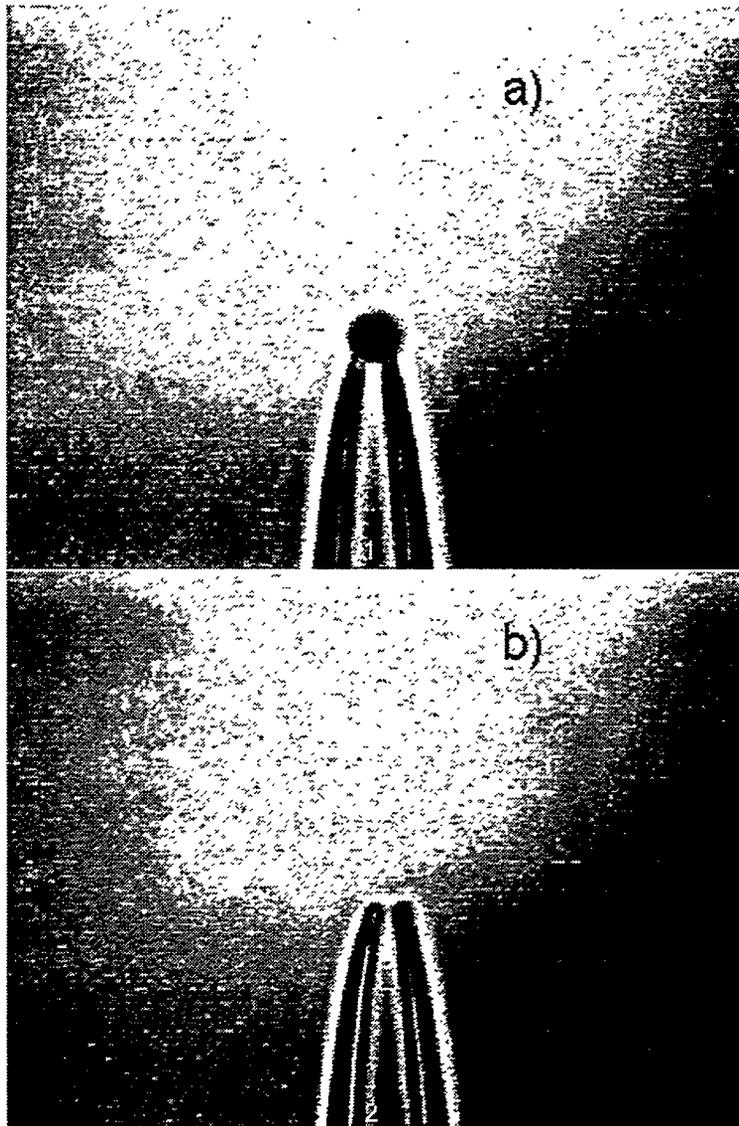


Figure 2:

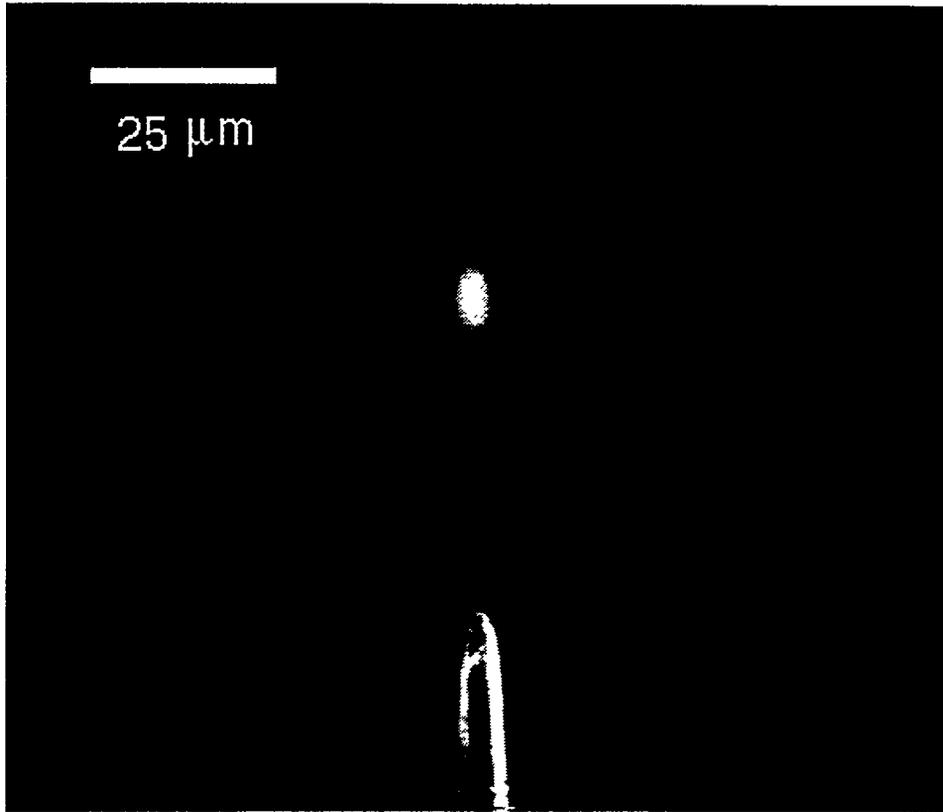


Figure 3:

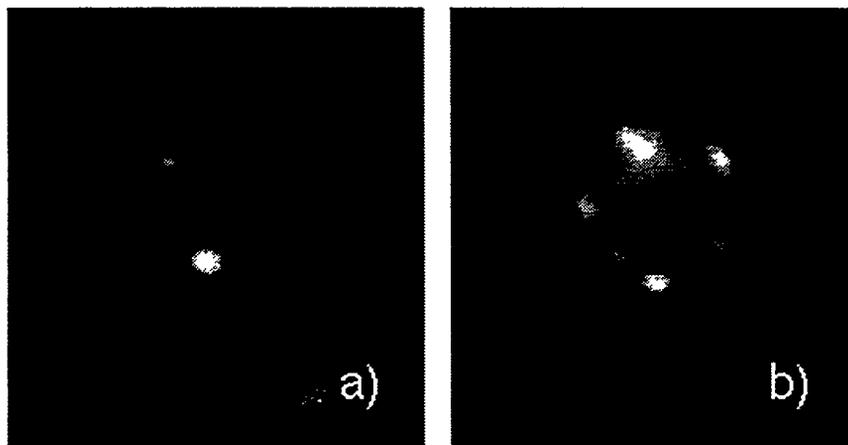


Figure 4:

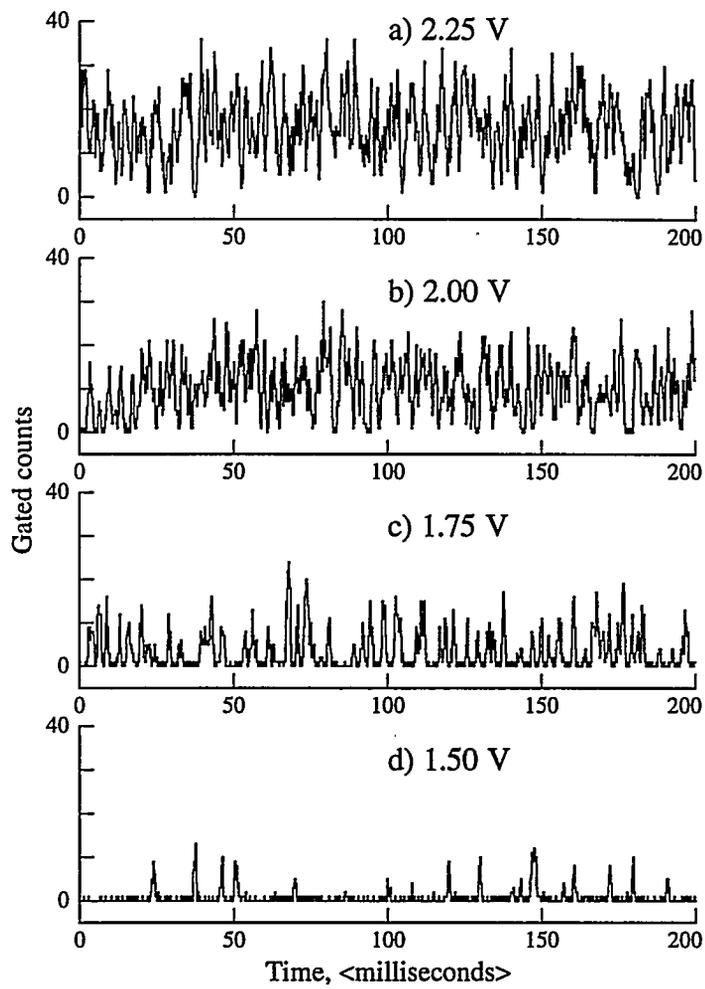


Figure 5:

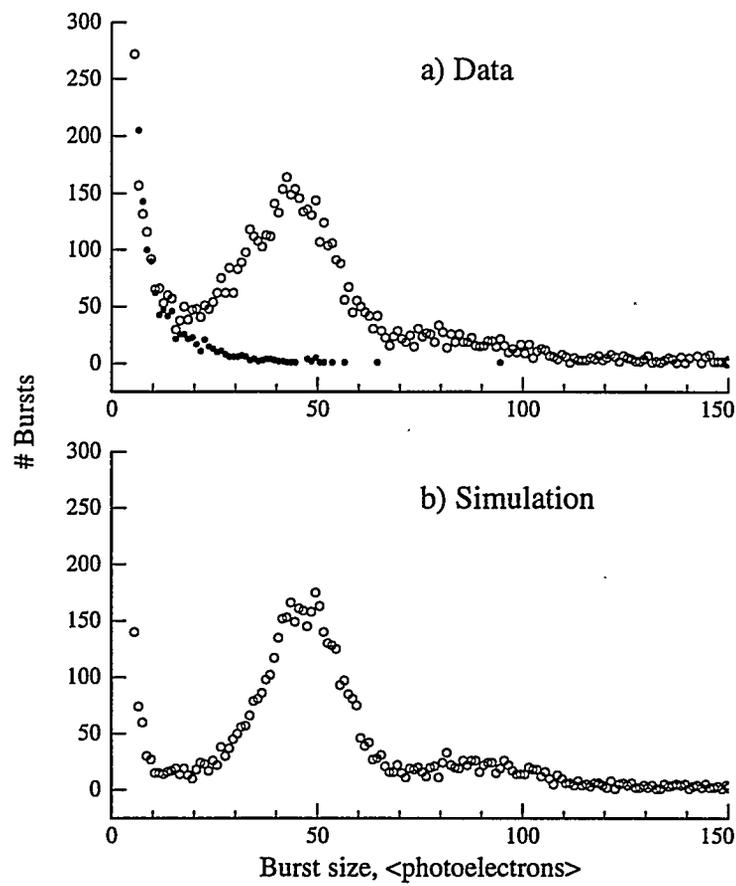


Figure 6:

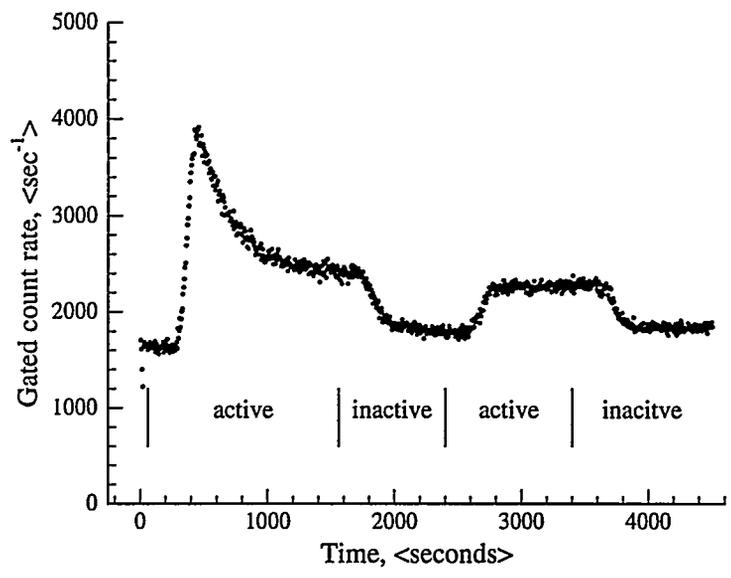


Figure 7:

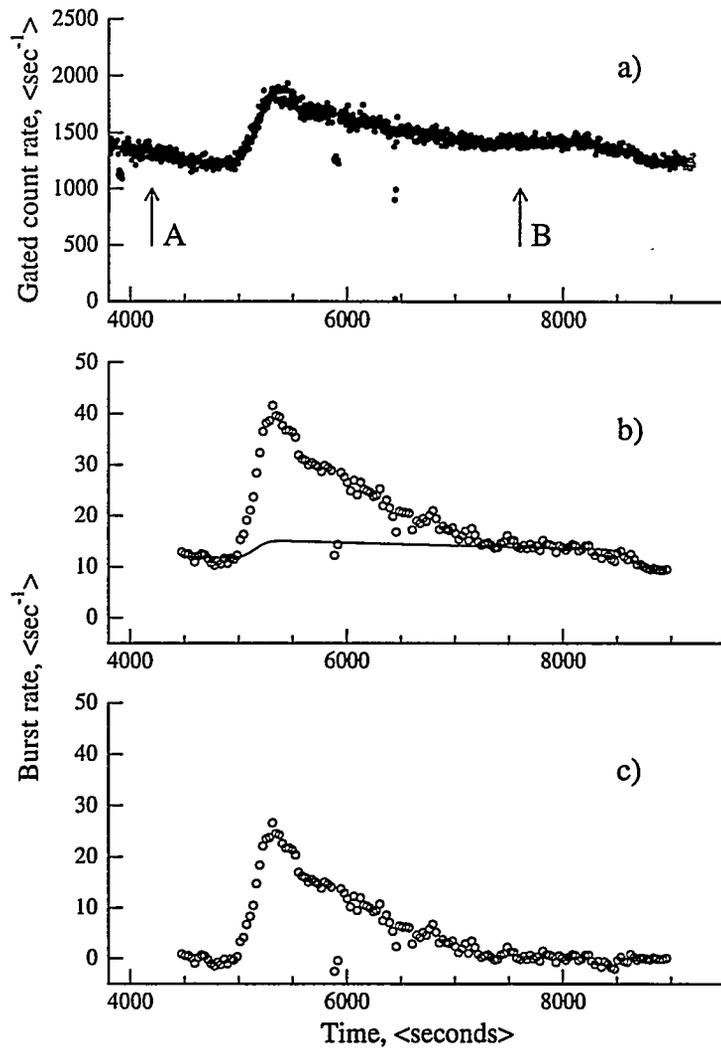


Figure 8: