

Single-molecule Fluorescence Spectroscopy of TOTO on Poly-AT and Poly-GC DNA[¶]

Benjamin P. Bowen^{*1}, Jörg Enderlein² and Neal W. Woodbury³

¹Department of Chemical and Materials Engineering, Arizona State University, Tempe, AZ;

²Institute for Biological Information Processing, Forschungszentrum Jülich, Jülich, Germany and

³Department of Chemistry and Biochemistry, Arizona Biodesign Institute, Arizona State University, Tempe, AZ

Received 17 July 2003; accepted 18 September 2003

ABSTRACT

Excited state lifetime and amplitude measurements were made on thiazole orange dimer (TOTO), a dimeric DNA-intercalating fluorophore, at single-molecule concentrations. As expected from previous study, the excited state lifetime of TOTO intercalated in DNA is dependent on the sequence of the double-stranded DNA, having values of 2.2 ns in poly-GC DNA and 1.8 ns in poly-AT DNA. The distribution of excited state lifetimes of single molecules of TOTO intercalated into oligonucleotides having varying proportions of poly-GC sequences relative to poly-AT sequences were analyzed as a function of the fraction of poly-GC. By using excited state lifetime distributions from the purely GC and purely AT oligonucleotides as a basis set, it was possible to estimate the GC content of oligonucleotides with intermediate GC composition to within a few percent error. This serves as a model for the analysis of equilibrium binding distributions in DNA using single-molecule methods.

INTRODUCTION

Homodimeric bis-intercalating fluorophores such as the thiazole orange dimer (TOTO) and the oxazole yellow dimer (YOYO) have been widely used for the study and identification of DNA (1). As is the case for other intercalating dyes, the quantum yield of fluorescence from TOTO increases dramatically (a factor of roughly 1400) when it is intercalated in double-stranded DNA (dsDNA) (1,2). In addition, TOTO binds strongly to dsDNA with a dissociation constant of approximately $1 \times 10^{-9} M$ (1,3). This dissociation constant is more than 100 times less than the dissociation constant for monomeric thiazole orange (TO). Nuclear magnetic resonance (NMR) structural studies have shown that four sequential nucleotides are involved in the TOTO–dsDNA interaction. The binding strength is thought to be sequence dependent

(4,5). For example, 5'-CTAG-3' has been identified as a preferred binding site in oligonucleotides containing naturally occurring DNA bases. Considerably stronger binding has been observed from certain sequences containing inosine or methyl-cytosine, which can improve binding strength by up to 68-fold over CTAG under conditions appropriate for NMR analysis (4). The techniques developed in this article are applied to study TOTO's site-specific binding affinity at the single-molecule level in the accompanying article under conditions more similar to those common during histochemical staining procedures (6).

Single-molecule fluorescence measurements of DNA-binding peptides attached covalently to the intercalating fluorophore TO have demonstrated that it is possible to detect and analyze fluorescence from single intercalating fluorophores (7). Previous measurements by Netzel *et al.* have shown that the excited state lifetimes of many intercalating fluorophores are multiexponential and dependent upon the composition of the DNA in which they are intercalated (2). Under the conditions used in that study, TOTO had an average excited state lifetime of 1.02 ns in ATAT regions of DNA and 2.03 ns in GCGC regions of DNA. A similar DNA sequence dependence of the excited state lifetime was also found for YOYO and many of the monomeric intercalating fluorophores. The sequence-dependent lifetime of intercalating dyes is potentially useful for analysis of GC content in stained samples. In a recent study by Murata *et al.* excited state lifetime images of single YOYO stained 3T3 cells were acquired to determine the AT and GC content of nuclear DNA during different phases of the cell cycle (8). Cells during different phases of growth showed excited state lifetimes ranging from 1.7 to 2.4 ns.

In recent years, considerable effort has been invested in the development of methods for identifying individual molecules on the basis of their fluorescence properties (9–11). The techniques used in this kind of study include pattern matching of the fluorescence decay and multidimensional correlation of fluorescence parameters. These methods rely on the ability to measure parameters accurately enough from each molecule such that different members of the population can be distinguished a large fraction of the time (12–14). This works well for molecules with strong single-molecule fluorescence signals relative to background noise. Unfortunately, the single-molecule fluorescence from intercalating dyes, similar to that from many single molecules of interest to biology, is much weaker than that from the best single-molecule fluorophores (6). This is likely because of the variety of photochemical reactions that the intercalating dye can undergo

[¶]Posted on the web site on 8 October 2003

*To whom correspondence should be addressed at: Department of Chemical and Materials Engineering, Arizona State University, Tempe, AZ 85287, USA. Fax: 480-965-2747; e-mail: bpb@asu.edu

Abbreviations: dsDNA, double-stranded DNA; IRF, instrumental response function; NMR, nuclear magnetic resonance; TCSPC, time-correlated single-photon counting; TO, thiazole orange; TOTO, thiazole orange dimer; YOYO, oxazole yellow dimer.

© 2003 American Society for Photobiology 0031-8655/03 \$5.00+0.00

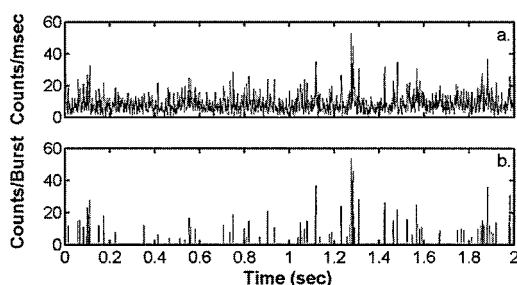


Figure 1. Fluorescence events from TOTO intercalated into the oligonucleotide GC8 (Table 1) are shown here. (a) The unmodified fluorescence data binned into millisecond time bins. (b) Results of the burst-finding algorithm for single-molecule events having three or more photons. See text for details.

otide to determine the bulk excited state lifetime of TOTO bound to poly-GC DNA. The bulk data was generated by merging all the fluorescence detected from single-molecules of the oligonucleotide in question. This was fit to a multiexponential fluorescence decay convolved with the instrumental response function (IRF) for the system. The fitting function included two exponential decay terms, a nondecaying component and a temporal shift in the IRF as free parameters. The shift parameter allows an accurate determination of the time zero and accounts for small timing shifts between the IRF and the fluorescence data. The value of the shift parameter is the difference between the time zero determined by the fit and the center of the original IRF. Figure 2 shows the IRF, raw data and fits of fluorescence from both poly-AT sequences (Fig. 2a) and poly-GC sequences (Fig. 2b) of DNA. Assuming that the nondecaying term in Fig. 2 is from background photons, nearly 100% of the initial fluorescence amplitude for the predominantly poly-AT sequence (GC8; Table 1) decays with a lifetime of 1.9 ns. Similarly, almost 100% of the fluorescence for the poly-GC sequence (GC100; Table 1) decays with a lifetime of 2.2 ns. In both cases, the remaining signal mostly consists of a small constant term. The contribution from subnanosecond fluorescence is only about 1% of the initial amplitude.

The distributions of measured excited state lifetimes from single molecules of TOTO intercalated in poly-AT DNA (GC8) and poly-GC DNA (GC100) are shown in Fig. 3a,b. Excited state lifetime distribution analysis occurred in two stages. First, the excited state lifetime distributions for the samples containing GC8 and GC100 were each fit to a Gaussian distribution resulting in a mean value, μ_1 , of 1.76 and a standard deviation, σ_1 , of 0.70 for poly-AT DNA, and a mean value, μ_2 , of 2.0 and a standard deviation, σ_2 , of 0.92 for poly-GC DNA. The mean and standard deviation from those distributions were used to describe the excited state lifetime distributions, $F(\tau)$, for samples with intermediate GC content (GC26, GC50 and GC76) according to the equation $F(\tau) = a_1G(\tau) + a_2H(\tau)$ where $G(\tau) = \exp[-(\tau - \mu_1)^2/2\sigma_1^2]$ and $H(\tau) = \exp[-(\tau - \mu_2)^2/2\sigma_2^2]$. The total integrated area under the distribution $F(\tau)$ is set to one; so that the values a_1 and a_2 are the relative amounts of AT and GC events in the dataset. Fitting of the distributions using the method described above was performed by least squares methods to determine the values of a_1 and a_2 . In Fig. 3f, the actual GC content vs the measured GC content (a_2) was fit to a linear curve, and the average residual error was 0.0226.

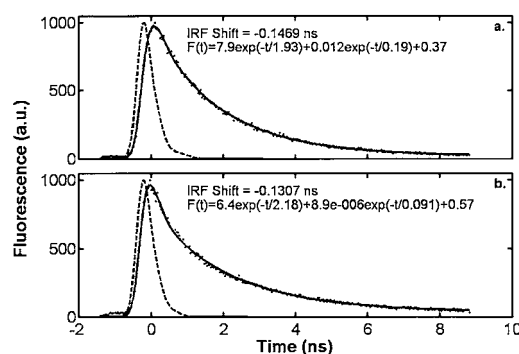


Figure 2. TCSPC decay curves for (a) GC8 and (b) GC100. Only photons identified as contributing to single-molecule bursts were used. Photons were divided into 35 ps time bins. The fitting curve convolved with a shifted IRF is shown as a solid line. The fitting functions with the coefficients resulting from the fit are shown. See text for details.

DISCUSSION

Excited state lifetime dependence on DNA sequence

In agreement with the study published by Netzel *et al.* regarding the excited state lifetime of TOTO in poly-AT and poly-GC regions of DNA, kinetic differences were observed for the two samples (2). However, the kinetics measured under the conditions described in this report differ from those measured by Netzel *et al.* The fluorescence of TOTO intercalated in poly-AT DNA is described by Netzel *et al.* using a three-component exponential decay having preexponential values and excited state lifetimes of (0.51, 0.17 ns); (0.33, 1.39 ns) and (0.16, 2.98 ns) respectively. The fluorescence of TOTO intercalated in poly-GC DNA is described by Netzel *et al.* using a three-component exponential decay having preexponential values and excited state lifetimes of (0.41, 0.21 ns); (0.29, 2.35 ns) and (0.30, 4.2 ns), respectively. In contrast, the kinetics observed in the current study are predominantly single exponential with lifetimes near 1.9 and 2.2 ns for poly-AT and poly-GC DNA sequences, respectively. These data were fit with two exponential decay terms and a constant term, but nearly 100% of the integrated fluorescence was accounted for in the primary decay component (Fig. 2). It is not clear why these results differ from Netzel *et al.*'s study. However, there are significant differences in experimental conditions, particularly regarding excitation wavelength. In this study, 495 nm laser pulses were used to excite the sample. This wavelength directly excites the lowest excited singlet state of the TOTO molecule. In the study reported by Netzel *et al.*, samples were excited by the third harmonic of the Nd³⁺yttrium–aluminum–garnet laser at 355 nm. Previous study has suggested that there is an excitation wavelength dependence of the fluorescence decay of YOYO, a related dimeric intercalating dye, bound to DNA (20).

Detection of single-molecules of TOTO on single oligonucleotides

Ideal single-molecule fluorescence probes have a large extinction coefficient, high quantum yield of fluorescence, low yield of dark state formation and high photostability. There has been little characterization of intercalating fluorophores at the single-molecule level, but bulk studies of TOTO intercalated in DNA have shown dynamic behavior such as single-stranded photocleavage and photobleaching that reduces the fluorescence signal (21,22). Thus

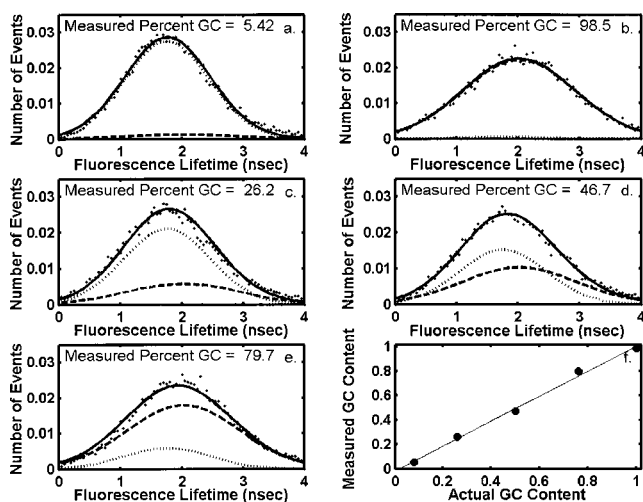


Figure 3. Distributions of excited state lifetimes and the associated Gaussian fits are shown for data collected from single molecules of TOTO intercalated into each of the oligonucleotides in Table 1: (a) GC8, (b) GC100, (c) GC26, (d) GC50 and (e) GC76. (f) The actual GC content vs the GC content determined from the lifetime distribution for each sample (panels a–e).

TOTO and DNA-intercalating dyes in general are far from ideal single-molecule probes in this respect, primarily because of their reactivity in the excited state. The single-molecule data presented here shows a factor of two to five times lower rate of fluorescence detectable from TOTO than one normally obtains with more robust dyes (23). The low quantum yield of fluorescence, which results in a large spacing between the detected photons, is likely because of a high quantum yield of dark state formation. However, intercalating dyes have some very interesting and useful properties as biological probes, particularly at the single-molecule level. First, they only fluoresce when incorporated into DNA. Second, some of them (TOTO being an example) tend to associate with DNA using the same kinds of interactions common for peptide or protein DNA ligands. In addition, intercalating dyes can be incorporated into DNA-binding proteins such that they only fluoresce when the protein binds to DNA, and the authors' laboratory has generated several such systems for both DNA-binding domains and for large complexes such as nucleosomes (7,15–17). Therefore, developing a single-molecule methodology for working with weakly emitting fluorophores (*i.e.* in this case, fluorophores whose fluorescent output saturates at relatively low laser intensity) or fluorophores with reactive excited states is of general interest for a number of biological applications.

Fluorescence data from 700 pM TOTO in solution with 100 nM dsDNA of sequence GC8 is shown in Fig. 1a,b. These concentrations ensure that all of the TOTO molecules are bound to DNA and that there are essentially no DNA molecules bound to multiple TOTO molecules. The detected photons observed as a function of time (in the lab time frame rather than relative to the mode-locked laser pulse) are binned with 1 ms long time intervals. Although there are some events yielding more than 30 photons/ms, the majority of events had roughly 10 photons/ms (the actual average burst size for these experiments was 13 photons).

The information available from any particular set of 13 photons is strongly limited by statistical noise. In this case, the primary

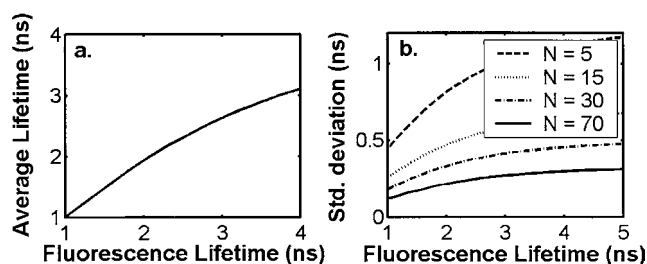


Figure 4. (a) The mean value of the photon arrival times as a function of the excited state lifetime according to Eq. 1. Note that for the range of lifetimes significant in this study (near 2 ns), the actual lifetime and the average arrival time are in good agreement. (b) The standard deviation (because of statistical error) for various numbers of photons, N , collected per burst according to Eq. 2.

parameter being considered is an excited state lifetime. The simplest way to generate an overall lifetime for a fluorescent decay is to average the arrival times (relative to the laser pulse) of all of the photons in an individual event. In the case of a true single exponential decay (and the TOTO decays here are predominantly one exponential), the average arrival time is in fact the excited state lifetime (even when the IRF has appreciable width assuming that it is symmetric).

The relationship between the actual excited state lifetime and the average photon arrival time can be determined as follows. Assuming a monoexponential fluorescence decay with decay time τ , the average arrival time of the detected photons would be

$$\langle t \rangle = \int_0^T t \cdot e^{-t/\tau} dt / \int_0^T e^{-t/\tau} dt \quad (1)$$

where T is the width of the TCSPC measurement window. One can show that for N photons per burst, the standard deviation (due to counting statistics) in the average arrival time is given by

$$\sigma_t^2 = \langle t^2 \rangle - \langle t \rangle^2 / N \quad (2)$$

where

$$\langle t^2 \rangle = \int_0^T t^2 \cdot e^{-t/\tau} dt / \int_0^T e^{-t/\tau} dt. \quad (3)$$

Figure 4a shows the theoretical relationship between the average arrival time of detected photons, $\langle t \rangle$, and the excited state lifetime, τ . The difference between $\langle t \rangle$ and τ becomes larger for longer fluorescence lifetimes because of the noninfinite measurement time window. This can also be seen in Fig. 4b, which shows the theoretical dependence of the standard deviation of the average arrival time on the fluorescent lifetime. Though a simple arrival time average does not yield an exact lifetime, particularly for lifetimes that are a substantial fraction of the time window of the measurement, it is a very rapid method for determining an effective lifetime of a large number of independent bursts and results in the lowest possible statistical noise. For the purposes of distinguishing between populations, minimum statistical noise is more important than accurate absolute lifetime.

The widths of the fluorescent lifetime distributions of Fig. 3 agree approximately with those predicted by Eq. 2 (Fig. 4b). As can be seen in Fig. 3, the width of each lifetime distribution is roughly 30% of the average lifetime value. Because of this wide

distribution, it would be impossible to reliably differentiate between individual TOTO molecules bound to a poly-AT region or a poly-GC region in a diffusing system such as this one, given only a 0.3 ns difference between them (the widths of the single-molecule lifetime distributions are about 0.6–0.7 ns; Fig. 3). However, even though single-molecule diffusion events cannot be identified individually, the distribution of excited state lifetimes from a large number of single-molecule measurements contains more information than does the ensemble lifetime of a bulk sample. An analysis of such distributions can be used to determine the equilibrium binding distribution of TOTO between different DNA sequences.

Population analysis of single-molecule fluorescence data

Even though the average lifetime difference between the two populations is small, the lifetime distributions are uniform, of known width, and are composed of only the two populations for the systems used here. Because the shapes of the distributions are critical in this analysis, large datasets are required. Each dataset contained approximately 16 750 events having a mean burst size of 13 photons.

The comparison of the predicted and actual GC content is plotted in Fig. 3f. There are two points worth discussing. First, the average residual error in the GC content determined from the distribution analysis is small, roughly 2% overall. Thus, even though the two populations have very similar lifetimes (1.8 vs 2.2 ns), and the error in measuring the lifetime of individual molecules is large (roughly ± 0.7 ns) it is possible to determine the relative fraction of each population in the distribution to a high degree of accuracy. Second, the fact that there is an absolute linear correspondence between the relative amount of the signal from GC vs AT sequences and the actual amount of GC vs AT sequence in the oligonucleotide implies that the binding constant of TOTO to GC and AT must be essentially identical (to within about 2% error).

An inherent limitation of this technique for determination of GC content in dsDNA fragments is that the fluorescence lifetime of TOTO presumably depends upon the four-base sequence surrounding the bis-intercalation site. In this study, dsDNA fragments containing continuous regions of poly-AT and poly-GC sequences were used. Whereas, in practice, an ideal technique for GC-content determination in dsDNA fragments would not require continuous regions of repetitive sequences similar to those used in this study.

Results from this study demonstrate the applicability of a method using single-molecule distributions of excited state lifetimes as a tool for analyzing two-state systems. In particular, this technique was used to measure the relative amounts of two subpopulations having only slightly different excited state lifetimes. TOTO–DNA interactions were used as a model system characteristic of situations in which the fluorophore under study has desirable biological properties but poor photophysical properties. This is possible because the population analysis method used is largely insensitive to the statistically limited width of the lifetime distribution of each subpopulation. Even though bursts with as few as three photons were included in the analysis, an accurate determination of the relative amounts of the two subpopulations was possible. Consequently, many systems that do not involve ideal fluorophores can be characterized quantitatively using single-molecule fluorescence methods that use distribution analysis.

Acknowledgements—This study was financially supported by NSF grant MCB0131766 and NSF grant DBI96022258003.

REFERENCES

- Rye, H. S., S. Yue, D. E. Wemmer, M. A. Quesada, R. P. Haugland, R. A. Mathies and A. N. Glazer (1992) Stable fluorescent complexes of double-stranded DNA with bis-intercalating asymmetric cyanine dyes—properties and applications. *Nucleic Acids Res.* **20**, 2803–2812.
- Netzel, T. L., K. Nafisi, M. Zhao, J. R. Lenhard and I. Johnson (1995) Base-content dependence of emission enhancements, quantum yields, and lifetimes for cyanine dyes bound to double-strand DNA: photophysical properties of monomeric and bichromophoric DNA stains. *J. Phys. Chem.* **99**, 17936–17947.
- Mecklenburg, M., A. Grauers, B. R. Jonsson, A. Weber and B. Danielsson (1997) A strategy for the broad range detection of compounds with affinity for nucleic acids. *Anal. Chim. Acta.* **347**, 79–86.
- Bunkenborg, J., M. M. Stidsen and J. P. Jacobsen (1999) On the sequence selective bis-intercalation of a homodimeric thiazole orange dye in DNA. *Bioconjugate Chem.* **10**, 824–831.
- Jacobsen, J. P., J. B. Pedersen, L. F. Hansen and D. E. Wemmer (1995) Site-selective bis-intercalation of a homodimeric thiazole orange-dye in DNA oligonucleotides. *Nucleic Acids Res.* **23**, 753–760.
- Bowen, B. P. and N. W. Woodbury (2003) TOTO binding affinity analysis using single-molecule fluorescence spectroscopy. *Photochem. Photobiol.* **78**, 579–583.
- Daniel, D. C., M. Thompson and N. W. Woodbury (2000) Fluorescence intensity fluctuations of individual labeled DNA fragments and a DNA binding protein in solution at the single molecule level: a comparison of photobleaching, diffusion, and binding dynamics. *J. Phys. Chem. B* **104**, 1382–1390.
- Murata, S., P. Herman and J. R. Lakowicz (2001) Texture analysis of fluorescence lifetime images of at- and gc-rich regions in nuclei. *J. Histochem. Cytochem.* **49**, 1443–1451.
- Eggeling, C., S. Berger, L. Brand, J. R. Fries, J. Schaffer, A. Volkmer and C. A. M. Seidel (2001) Data registration and selective single-molecule analysis using multi-parameter fluorescence detection. *J. Biotechnol.* **86**, 163–180.
- Fries, J. R., L. Brand, C. Eggeling, M. Kollner and C. A. M. Seidel (1998) Quantitative identification of different single molecules by selective time-resolved confocal fluorescence spectroscopy. *J. Phys. Chem. A* **102**, 6601–6613.
- Zander, C., M. Sauer, K. H. Drexhage, D. S. Ko, A. Schulz, J. Wolfrum, L. Brand, C. Eggeling and C. A. M. Seidel (1996) Detection and characterization of single molecules in aqueous solution. *Appl. Phys. B-Lasers Opt.* **63**, 517–523.
- Enderlein, J. and M. Sauer (2001) Optimal algorithm for single-molecule identification with time-correlated single-photon counting. *J. Phys. Chem. A* **105**, 48–53.
- Sauer, M., J. Arden-Jacob, K. H. Drexhage, F. Gobel, U. Lieberwirth, K. Muhlegger, R. Muller, J. Wolfrum and C. Zander (1998) Time-resolved identification of individual mononucleotide molecules in aqueous solution with pulsed semiconductor lasers. *Bioimaging* **6**, 14–24.
- Herten, D. P., P. Tinnefeld and M. Sauer (2000) Identification of single fluorescently labelled mononucleotide molecules in solution by spectrally resolved time-correlated single-photon counting. *Appl. Phys. B-Lasers Opt.* **71**, 765–771.
- Thompson, M. and N. W. Woodbury (2001) Thermodynamics of specific and nonspecific DNA binding by two DNA-binding domains conjugated to fluorescent probes. *Biophys. J.* **81**, 1793–1804.
- Thompson, M. and N. W. Woodbury (2000) Fluorescent and photochemical properties of a single zinc finger conjugated to a fluorescent DNA-binding probe. *Biochemistry* **39**, 4327–4338.
- Babendure, J., P. A. Liddell, R. Bash, D. LoVullo, T. K. Schiefer, M. Williams, D. C. Daniel, M. Thompson, A. K. Taguchi, D. Lohr and N. W. Woodbury (2003) Development of a fluorescent probe for the study of nucleosome assembly and dynamics. *Anal. Biochem.* **317**, 1–11.
- Bowen, B. and N. Woodbury (2003) Single-molecule fluorescence lifetime and anisotropy measurements of the red fluorescent protein, dsred, in solution. *Photochem. Photobiol.* **77**, 362–369.

19. Enderlein, J., D. L. Robbins, W. P. Ambrose and R. A. Keller (1998) Molecular shot noise, burst size distribution, and single-molecule detection in fluid flow: effects of multiple occupancy. *J. Phys. Chem. A* **102**, 6089–6094.
20. Abramo, K. H., J. B. Pitner and L. B. McGown (1998) Spectroscopic studies of single-stranded DNA ligands and oxazole yellow dyes. *Biospectroscopy* **4**, 27–35.
21. Akerman, B. and E. Tuite (1996) Single- and double-strand photocleavage of DNA by yo, yoyo and toto. *Nucleic Acids Res.* **24**, 1080–1090.
22. Kanony, C., B. Akerman and E. Tuite (2001) Photobleaching of asymmetric cyanines used for fluorescence imaging of single DNA molecules. *J. Am. Chem. Soc.* **123**, 7985–7995.
23. Wennmalm, S. and R. Rigler (1999) On death numbers and survival times of single dye molecules. *J. Phys. Chem. B* **103**, 2516–2519.