Anal. Chem. **2006,** *78,* 2039−2050

Single-Molecule Detection and Identification of Multiple Species by Multiparameter Fluorescence Detection

Jerker Widengren,*,†,‡ Volodymyr Kudryavtsev,‡,§ Matthew Antonik,§ Sylvia Berger,§,∥ Margarita Gerken,†,⊥ and Claus A. M. Seidel*,§

Department of Applied Physics, Royal Institute of Technology, Albanova University Center, 10691 Stockholm, Sweden, and Lehrstuhl für Molekulare Physikalische Chemie, Heinrich-Heine Universität Düsseldorf, Universitätsstrasse 1, D-40225 Düsseldorf, Germany

Two general strategies are introduced to identify and quantify single molecules in dilute solutions by employing a spectroscopic method for data registration and specific burst analysis, denoted multiparameter fluorescence detection (MFD). MFD uses pulsed excitation and time-correlated single-photon counting to simultaneously monitor the evolution of the eight-dimensional fluorescence information (fundamental anisotropy, fluorescence lifetime, fluorescence intensity, time, excitation spectrum, fluorescence spectrum, fluorescence quantum yield, distance between fluorophores) in real time and allows for selection of specific events for subsequent analysis. Using the multiple fluorescence dimensions, we demonstrate a dye labeling scheme of oligonucleotides, by which it is possible to identify and separate 16 different compounds in the mixture via their characteristic pattern by MFD. Such identification procedures and multiplex assays with single-molecule sensitivity may have a great impact on screening of species and events that do not lend themselves so easily to amplification, such as disease-specific proteins and their interactions.

Assays enabling detection and identification of minute amounts of biomolecules have attracted a strong interest in the last years. In the biomedical field, the ability to detect and identify disease-specific proteins in low numbers, specific for an incipient cancerous disease or a bacterial or viral infection, can provide detailed early-stage diagnoses. Likewise, technologies for sensitive, precise, and accurate detection of the genetic bases underlying onset of diseases and variations in drug responses (pharmacogenetics) and exposure to toxic agents (toxicogenetics) are likely to have a dramatic impact on future diagnostics and (individualized) pharmacological treatments.[1]

Such pharmacogenetical and toxicogenetical differences can often be attributed to single base-pair variations in the genome, so-called single-nucleotide polymorphisms, which are of particular interest to detect and identify.[2] Over the past decade, fluorescence-based techniques for molecular studies have shown remarkable progress. Fluorescence-based platforms have been developed for sensitive and high-throughput detection and identification of nucleic acids, proteins, and other biomolecules.[3] These platforms often take advantage of the specificity of naturally occurring enzymes and recognition molecules, such as in Taqman genotyping. By using molecular beacons, aptamers, and monoclonal antibodies, the molecules searched for can be presented to arrays of target molecules in a highly parallelized manner.[4]

The exquisite sensitivity of fluorescence detection has made detection, identification, and characterization of single molecules possible. This not only enables fundamental scientific studies of dynamic and conformational properties of single biomolecules,[5] it also provides a potential to establish molecular profiles, with an ultimate level of sensitivity, of species present at very low concentrations or quantities.[6−9]

For nucleic acid analysis, single-molecule techniques make it possible to circumvent amplification steps, thereby avoiding quantification difficulties due to small differences in efficiencies of amplifications from sample to sample and from one thermal cycle to another, as experienced, for example, for PCR-based assays. Several fluorescence-based approaches have also been reported for single-molecule DNA sequencing.[9−16] However, single-molecule sensitivity may have an even greater impact on

* Corresponding authors. E-mail: jerker@biomolphysics.kth.se.
† Royal Institute of Technology.
‡ Contributed equally.
§ Heinrich-Heine Universität Düsseldorf.
∥ Present address: Fluka Chemie GmbH, Industriestrasse 25, CH-9471 Buchs, Switzerland.
⊥ Present address: 3. Physikalisches Institut, Pfaffenwaldring 57, D-70550 Stuttgart, Germany.

(1) Ross, J. S.; Ginsburg, G. S. *DDT* **2002**, *7*, 859−64.
(2) Pfost, D. R.; Boyce-Jacino, M. T.; Grant, D. M. *Trends Biotechnol.* **2000**, *18*, 334−8.
(3) Eggeling, C.; Brand, L.; Ullmann, D.; Jager, S. *DDT* **2003**, *8*, 632−41.
(4) Shi, M. M. *Clin. Chem.* **2001**, *47*, 164−72.
(5) Moerner, W. E.; Fromm, D. P. *Rev. Sci. Instrum.* **2003**, *74*, 3597−619.
(6) Knemeyer, J.-P.; Marmé, N.; Sauer, M. *Anal. Chem.* **2000**, *72*, 3717−24.
(7) Anazawa, T.; Matsunaga, H.; Yeung, E. S. *Anal. Chem.* **2002**, *74*, 5033−8.
(8) Herten, D. P.; Tinnefeld, P.; Sauer, M. *Appl. Phys. B* **2000**, *71*, 765−71.
(9) Tinnefeld, P.; Sauer, M. *Angew. Chem., Int. Ed.* **2005**, *44*, 2642−71.
(10) Neuweiler, H.; Sauer, M. *Anal. Chem.* **2005**, *77*, 178A−85A.
(11) Scheffler, S.; Sauer, M.; Neuweiler, H. *Z. Phys. Chem.* **2005**, *219*, 665−78.
(12) Jett, J. H.; Keller, R. A.; Martin, J. C.; Marrone, B. L.; Moyzis, R. K.; Ratliff, R. L.; Seitzinger, N. K.; Shera, E. B.; Stewart, C. C. J. *J. Biomol. Struct. Dyn.* **1989**, *7*, 301−9.
(13) Werner, J. H.; Cai, H.; Jett, J. H.; Reha-Krantz, L.; Keller, R. A.; Goodwin, P. M. *J. Biotechnol.* **2003**, *102*, 1−14.

10.1021/ac0522759 CCC: $33.50
Published on Web 02/07/2006

the screening of species and events that do not lend themselves so easily to amplification, such as disease-specific proteins and their interactions.[11,17]

In recent years, the characterization of molecules in single-molecule fluorescence detection measurements has step by step been established for the different fluorescence parameters. Spectral properties of absorption and fluorescence, $F(\lambda_A, \lambda_F)$,[18] fluorescence brightness and quantum yield, Φ_F,[19] fluorescence lifetime, τ,[20−26] and anisotropy, r,[27,28] are the five intrinsic properties of a fluorophore that are accessible in a multiparameter fluorescence detection (MFD) experiment (see Figure 1). These "chromophore parameters" can be deduced from the time-resolved detection of the five observables of the chromophore, which now serves as a tool to report on its local environment.

Considering molecular systems, changes in fluorescence parameters of a single coupled fluorophore sometimes do not provide enough information for molecular identification or for more detailed investigations of molecular interactions. Further information can be obtained on a single-molecule level by having more than one fluorophore per particle involved, thereby increasing the possibilities to determine stoichiometries and interactions of individual particles from photon densities and coincidences.[29,30] Moreover, the use of two fluorescing reporters also makes it possible to determine structural features of particles via Förster fluorescence resonance energy transfer (FRET).

In FRET, the energy from an excited donor fluorophore, D, is nonradiatively transferred to an acceptor fluorophore, A, by a strongly distance-dependent dipole−dipole coupling. Via the measured FRET efficiency, long-range molecular distance information can be provided, in a range of 20−100 Å, which is not covered by virtually any other solution technique.[31−33] FRET is

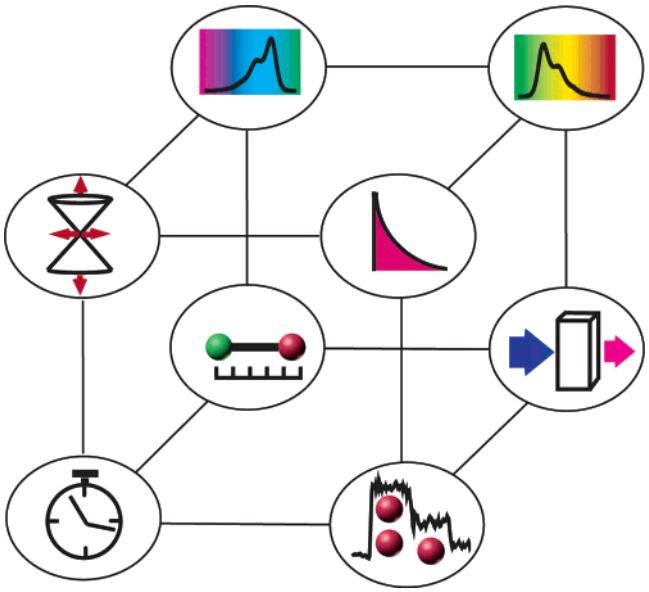

Figure 1. Multiple dimensions of the fluorescence. Sketches for fluorescence properties at the front side of the cube (starting at the upper left corner with clockwise rotation): fundamental anisotropy, fluorescence lifetime, fluorescence intensity, time; backside: excitation spectrum, fluorescence spectrum, fluorescence quantum yield, and distance between fluorophores.

taken advantage of in a number of screening procedures, providing information about changes in inter- and intramolecular distances that take place as an effect of the specific molecular interaction that is screened for.[34] When performed at the single-molecule level, FRET studies can yield information about heterogeneities in terms of conformations and conformational dynamics that are unavailable from ensemble measurements.[35] With a donor−acceptor distance approaching or exceeding 100 Å, the FRET efficiency typically is so small that the fluorescence from the acceptor can be determined only by additional direct excitation such as two-photon[36] or two-color excitation in a cw or alternating mode (ALEX).[37] Then the simultaneous presence of the two dyes, significant for an interaction or binding event, can be determined by cross-correlation,[36] colocalization,[38] or two-dimensional fluorescence intensity distribution analysis.[39]

Following the ideas of Förster, it is obvious that the fluorescence has multiple characteristic parameters that can be exploited to yield specific identification by a multiplex analysis.[40−43] In view

(14) Sauer, M.; Angerer, B.; Ankenbauer, W.; Földes-Papp, Z.; Gobel, F.; Han, K. T.; Rigler, R.; Schulz, A.; Wolfrum, J.; Zander, C. *J. Biotechnol.* **2001**, *86*, 181−201.
(15) Dörre, K.; Brakmann, S.; Brinkmeier, M.; Han, K. T.; Riebeseel, K.; Schwille, P.; Stephan, J.; Wetzel, T.; Lapczyna, M.; Stuke, M.; Bader, R.; Hinz, M.; Seliger, H.; Holm, J.; Eigen, M.; Rigler, R. *Bioimaging* **1997**, *5*, 139−52.
(16) Braslavsky, I.; Hebert, B.; Kartalov, E.; Quake, S. R. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 3960−4.
(17) Sauer, M.; Zander, C.; Müller, R.; Ullrich, B.; Drexhage, K. H.; Kaul, S.; Wolfrum, J. *Appl. Phys.* **1997**, *65*, 427−31.
(18) Tamarat, P.; Maali, A.; Lounis, B.; Orrit, M. *J. Phys. Chem.* **2000**, *104*, 1−16.
(19) Kask, P.; Palo, K.; Ullmann, D.; Gall, K. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 13756−61.
(20) Zander, C.; Sauer, M.; Drexhage, K. H.; Ko, D. S.; Schulz, A.; Wolfrum, J.; Brand, L.; Eggeling, C.; Seidel, C. A. M. *Appl. Phys. B* **1996**, *63*, 517−23.
(21) Edman, L.; Mets, Ü.; Rigler, R. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 6710−5.
(22) Maus, M.; Cotlet, M.; Hofkens, J.; Gensch, T.; De Schryver, F. C.; Schaffer, J.; Seidel, C. A. M. *Anal. Chem.* **2001**, *73*, 2078−86.
(23) Bowen, B. P.; Scruggs, A.; Enderlein, J.; Sauer, M.; Woodbury, N. *J. Phys. Chem. A* **2004**, *108*, 4799−804.
(24) Enderlein, J.; Sauer, M. *J. Phys. Chem. A* **2001**, *105*, 48−53.
(25) Tellinghuisen, J.; Goodwin, P. M.; Ambrose, W. P.; Martin, J. C.; Keller, R. A. *Anal. Chem.* **1994**, *66*, 64−72.
(26) Enderlein, J.; Goodwin, P. M.; van Orden, A.; Ambrose, W. P.; Erdmann, R.; Keller, R. A. *Chem. Phys. Lett.* **1997**, *270*, 464−70.
(27) Ha, T.; Laurence, T. A.; Chemla, D. S.; Weiss, S. *J. Phys. Chem. B* **1999**, *103*, 6839−50.
(28) Schaffer, J.; Volkmer, A.; Eggeling, C.; Subramaniam, V.; Striker, G.; Seidel, C. A. M. *J. Phys. Chem. A* **1999**, *103*, 331−6.
(29) Schmidt, T.; Schütz, G. J.; Gruber, H. J.; Schindler, H. *Anal. Chem.* **1996**, *68*, 4397−401.
(30) Weston, K. D.; Dyck, M.; Tinnefeld, P.; Muller, C.; Herten, D. P.; Sauer, M. *Anal. Chem.* **2002**, *74*, 5342−9.
(31) Lilley, D. M. J.; Wilson, T. J. *Curr. Opin. Chem. Biol.* **2000**, *4*, 507−17.
(32) Selvin, P. R. *Nat. Struct. Biol.* **2000**, *7*, 730−4.
(33) van der Meer, B. W.; Cooker, G.; Chen, S. Y. *Resonance Energy Transfer: Theory and Data*; VCH Publishers: New York, 1994.
(34) Karlstrom, A.; Nygren, P. A. *Anal. Biochem.* **2001**, *295*, 22−30.
(35) Ha, T. *Methods* **2001**, *25*, 78−86.
(36) Heinze, K. G.; Koltermann, A.; Schwille, P. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 10377−82.
(37) Kapanidis, A. N.; Laurence, T. A.; Lee, N. K.; Margeat, E.; Kong, X. X.; Weiss, S. *Acc. Chem. Res.* **2005**, *38*, 523−33.
(38) Lacoste, T. D.; Michalet, X.; Pinaud, F.; Chemla, D. S.; Alivisatos, A. P.; Weiss, S. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 9461−6.
(39) Kask, P.; Palo, K.; Fay, N.; Brand, L.; Mets, Ü.; Ullmann, D.; Jungmann, J.; Pschorr, J.; Gall, K. *Biophys. J.* **2000**, *78*, 1703−13.
(40) Förster, T. *Fluoreszenz organischer Verbindungen*, Vandenhoeck & Ruprecht: Göttingen, Germany, 1951.
(41) Förster, T. *Ann. Phys.* **1948**, *2*, 55−75.
(42) Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*, 2nd ed.; Kluwer Academic/Plenum Publishers: New York, 1999.
(43) Valeur, B. *Molecular Fluorescence: Principles and Applications*; Wiley-VCH Verlag: Weinheim, Germany. 2002.

of the additional information that can be extracted from a molecule having two or more fluorophores coupled to it, it is reasonable to add additional fluorescence dimensions to the five "chromophore parameters" stated above. With reference to Figure 1, we argue that the fluorescence parameter space can under appropriate conditions be considered to be at least eight-dimensional, adding three "system/environmental parameters" [(fluorophore) stoichiometry, (fluorophore) coupling, and time (at time scales other than that of excitation and emission to and from the excited singlet state)] to the above five "chromophore" parameters. It is noteworthy, that the two dyes can interfere with the fluorescence of one another by time-dependent intrinsic (e.g., transitions to and from photophysical states other than the singlet ground and excited states) or extrinsic (e.g., quenching or other chemical reactions influencing the fluorescence) processes.

Given that the FRET efficiency is reflected in several of the fluorescence parameters of the donor and acceptor fluorophores, a multidimensional and simultaneous registration of the fluorescence parameters from both the fluorophores can therefore increase both precision and accuracy. Moreover, since local effects also influence the fluorescence properties of the fluorophores, a multidimensional registration of the fluorescence can increase the certainty with which differences in the measured parameters due to true changes in the FRET efficiency and the interdye distances can be distinguished from differences due to local effects.[44,45] MFD has recently been successfully applied at the single-molecule level to study heterogeneities in conformations and conformational dynamics of HIV-1 reverse transcriptase:primer/template complexes[46] as well as to reveal a dynamic equilibrium between closed and open forms of the SNARE protein syntaxin-1, with possible implications for the regulation of membrane fusion of synaptic vesicles.[47]

The aim of this study is not to deliver a detailed analysis of specific molecular dynamic processes. Rather, it presents an investigation to what extent FRET-active compounds can be identified on the single-molecule level utilizing MFD and the differences in the detected fluorescence parameters of the molecules. It has previously been shown that single fluorescently labeled mononucleotides, with a specific fluorophore labeled to each of four different types of mononucleotides, can be detected and identified using spectrally resolved time-correlated single-photon counting.[48]

Here, we combine single-molecule identification based on fluorophore label identity with identification based on differences in intramolecular distances (donor-to-acceptor distances). Single molecules are analyzed, as they are freely diffusing in to and out of a confocal detection volume in a homogeneous sample. This approach offers both advantages and disadvantages, compared to the detection of surface-immobilized molecules. On one hand, the investigated molecules undertake different trajectories through the detection volume, which leads to a distribution of both the experienced excitation intensities and the interrogation times. This generates a variation of the brightness and duration of the fluorescence bursts, and many fluorescence bursts are too small to unequivocally be identified as a single-molecule event. MFD, including burst detection and subsequent offline data analysis, is useful in obtaining a sample survey, based only on those single-molecule events that have a long pathway through the detection volume.[49] The information content, in terms of the number of fluorescence photons detected per molecule, is limited by the dwell times, which typically cannot be extended maximally, until the fluorophore marker(s) eventually is (are) bleached. The great advantage of this homogeneous assay format is very good statistics on the investigated sample with several thousands of analyzed molecules, which is easily obtained. Moreover, for surface-immobilized molecules, evidently no information on the diffusion properties of the investigated molecules can be obtained. In addition, in an inhomogeneous assay, molecular and fluorescence marker properties can be significantly influenced by the immobilization itself. Immobilization also does not fulfill the demand for single-molecule detection and identification of molecules without the need for prior separation or manipulation steps. However, the MFD approach presented here is also fully employable and useful for a wide range of molecules, regardless whether immobilization is used or not.

In this study, 12 identical oligonucleotides, differing in the combination of dyes labeled to them, but also in the distance between the two dyes, were measured separately at the single-molecule level. To simulate a mixture measurement, the results of the burstwise analysis, which contain the fluorescence parameters for each molecule of all 12 species, were combined into a single data set. This postmeasurement mixing is possible because, in single-molecule measurements, the burst events are independent of each other, and the order of the events plays no role in the analysis. We introduce two different approaches for how individual molecules among samples of 12 differently labeled oligonucleotides can be identified. They both rely on sequential separations in different two-dimensional parameter space projections, or along several one-dimensional parameter axes.

In the first approach, a graphical analysis is performed to display the MFD features visually. The starting point is to arrange all detected molecules into a series of multidimensional histograms, based on the measured fluorescence parameters from each individual molecule. Within the total collection of detected molecules, and from the differences in the fluorescence parameters detected among the molecules, we show how it is possible to separate and identify the different types of oligonucleotides. Also shown is the possibility to discriminate against impurities, free dyes, and oligonucleotides labeled only with a single donor dye, yielding all in all 16 different compounds that are identified.

In the second approach, a mathematical procedure is applied that converts the above stepwise graphical analysis into a one-step computer algorithm. We introduce an identification algorithm that compares the measured fluorescence parameters for each detected molecule with the probability distributions for the

(44) Eggeling, C.; Berger, S.; Brand, L.; Fries, J. R.; Schaffer, J.; Volkmer, A.; Seidel, C. A. M. *J. Biotechnol.* **2001**, *86*, 163−80.
(45) Kühnemuth, R.; Seidel, C. A. M. *Single Mol.* **2001**, *2*, 251−4.
(46) Rothwell, P. J.; Berger, S.; Kensch, O.; Felekyan, S.; Antonik, M.; Wöhrl, B. M.; Restle, T.; Goody, R. S.; Seidel, C. A. M. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 1655−60.
(47) Margittai, M.; Widengren, J.; Schweinberger, E.; Schröder, G. F.; Felekyan, S.; Haustein, E.; König, M.; Fasshauer, D.; Grubmüller, H.; Jahn, R.; Seidel, C. A. M. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 15516−21.
(48) Lieberwirth, U.; Arden-Jacob, J.; Drexhage, K. H.; Herten, D. P.; Müller, R.; Neumann, M.; Schulz, A.; Siebert, S.; Sagner, G.; Klingel, S.; Sauer, M.; Wolfrum, J. *Anal. Chem.* **1998**, *70*, 4771−9.
(49) Fries, J. R.; Brand, L.; Eggeling, C.; Köllner, M.; Seidel, C. A. M. *J. Phys. Chem. A* **1998**, *102*, 6601−13.

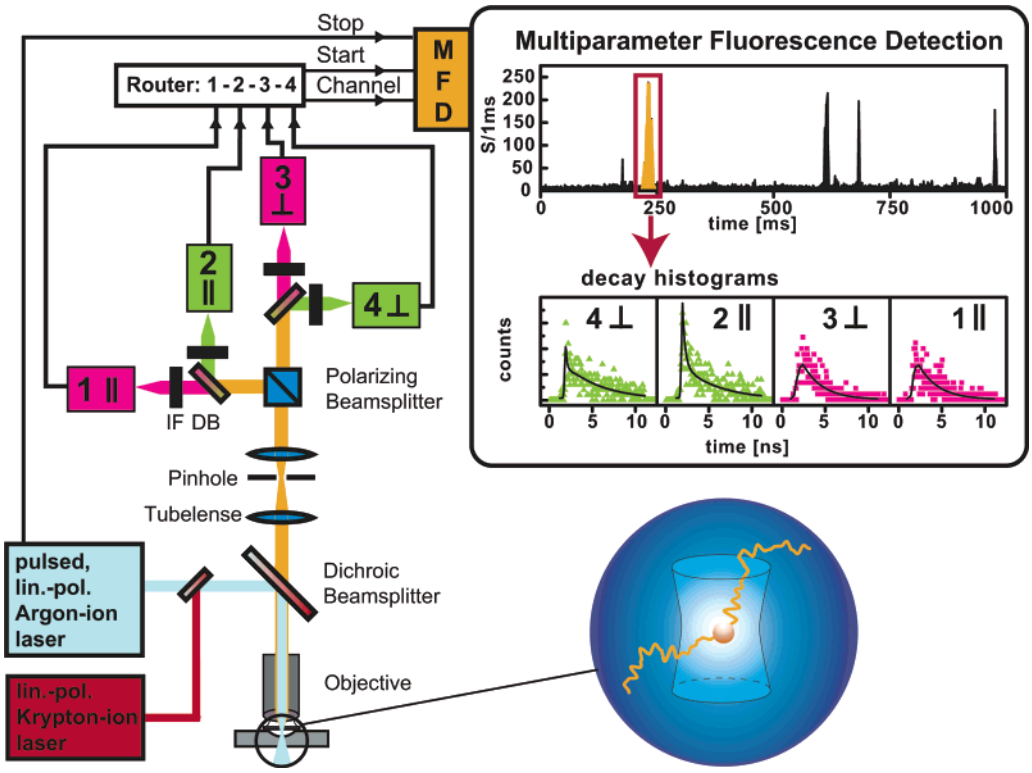

Figure 2. Four-channel confocal setup for simultaneous detection of spectral range, λ_F fluorescence intensity S, lifetime τ, anisotropy r. Abbreviations: IF, interference filter; DB, dichroic beam splitter.

corresponding parameters for all different possible species. The comparison is performed for several independent parameters or parameter combinations. The approach can be shown to offer fast, stable, and reliable species identification and to determine concentration profiles of molecular species with high fidelity.

Based on these results, we suggest confocal single-molecule spectroscopy using MFD as a useful tool for qualitative analysis revealing the presence of multiple distinct molecules in a sample. This study demonstrates the precision and accuracy with which single molecules can be detected, identified, and analyzed and with which concentration profiles of multiple molecular species can be determined, with possible implications for diagnostics of profiles of different molecules, their presence and interactions.[9−11]

METHODS AND MATERIALS

Setup. The confocal microscopy arrangement used in this study is depicted in Figure 2. The setup has basic features similar to those previously presented for applications of ultrasensitive fluorescence spectroscopy.[50,51] Freely diffusing molecules are excited by a pulsed, linearly polarized argon ion laser emitting at 514 nm (Sabre, Coherent, Palo Alto, CA) with a repetition rate of 73 MHz and an excitation power of 558 μW. In addition, there is the option to use a linear polarized krypton ion laser emitting at 647 nm (Sabre, Coherent) for two-color excitation experiments. The two lasers are not yet synchronized. After reflection by a dual-band dichroic mirror (DCLP510, AHF Analysentechnik), the laser beam is focused by a microscope objective (Olympus 60×, NA 1.2, water immersion). The fluorescence is collected and then refocused by the same objective to the image plane, where a pinhole is placed (radius 100 μm). From a FCS diffusion time of free rhodamine 6G (Rh6G) dye of τ_D = 256 μs, the detection volume can be estimated to be ~1 fL. After passage through the pinhole, the fluorescence is separated by a polarizing beam splitter, based on its direction of polarization with respect to that of the laser beam. The fluorescence is then divided by a dichroic beam splitter (510DCLX, AHF Analyzentechnik) into the fractions originating from the acceptor [cyanine 5 (Cy5) or BODIPY650 (BP650/665)] and donor (cyanine 3 (Cy3) or Rh6G fluorophores). Subsequently, band-pass filters are used (HQ730/140 by AHF Analyzentechnik for Cy5 and BP650/665; HQ575/70 by AHF Analyzentechnik for Cy3, Rh6G) to discriminate the two fractions of the collected fluorescence from scattered laser light. Each of the fractions of the fluorescence light is then detected by an avalanche photodiode (SPCM AQ-141, EG&G Vaudreuil, Quebec, Canada), the output pulses of which are processed online by a PC-BIFL-card with a routing unit (SPC 432 Becker&Hickl GmbH, Berlin, Germany) in the MFD mode.

MFD. The data registration and processing of MFD has been described in detail.[44] In brief, using the setup depicted in Figure 2, each detected photon can be characterized by four parameters: (i) spectral range, λ_F, of the detected fluorescence (green or red); (ii) the polarization of the signal photon (parallel or perpendicular) with respect to the linear polarization of the excitation laser; (iii) the arrival time of the signal photon relative to the incident laser pulse; and (iv) the interphoton time, Δt, to the preceding signal photon.

(50) Rigler, R.; Mets, Ü.; Widengren, J.; Kask, P. *Eur. Biophys. J.* **1993**, *22*, 169−75.
(51) Eggeling, C.; Fries, J. R.; Brand, L.; Günther, R.; Seidel, C. A. M. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 1556−61.

A single molecule diffusing through the laser focus generates a burst of fluorescence photons. Thus, a single-molecule event (burst) can be separated from background via small interphoton times. In addition to a criterion based on a maximum interphoton time of 0.0493 ms, bursts also had to contain a minimum total number of detected photons, here set to 120, to be registered as a molecular event. Once a single-molecule event is identified, average green and red count rates of the burst, S_G and S_R, are calculated by dividing the number of registered photons by the individual burst duration times. Arrival times of the photons in the selected burst are used to generate fluorescence decay histograms for each of the four detectors (green parallel, $G_{||}$, green perpendicular, $G_{\perp}$, red parallel, $R_{||}$, red perpendicular, $R_{\perp}$). With further analysis, the parameters lifetime, τ, and anisotropy, r, for each spectral range are determined.[28] The anisotropy was primarily analyzed in the green range. The steady-state anisotropy is formally given by

$$r_G = \int_0^\infty r_G(t)F_G(t)\ dt / \int_0^\infty F_G(t)\ dt$$

where

$$r_G(t) = \frac{F_{G||}(t) - F_{G\perp}(t)}{F_{G||}(t) + 2F_{G\perp}(t)}$$

is the time-dependent anisotropy, and where $F_{G||}(t)$ and $F_{G\perp}(t)$ are the donor fluorescence detected with a polarization parallel and perpendicular to that of the excitation light, respectively. $F_G(t) = F_{G||}(t) + 2F_{G\perp}(t)$ is the total detected donor fluorescence intensity. In this study, the anisotropy of the fluorescence detected from a single molecule in the green channels was determined as by Schaffer et al.,[28] from

$$r_G = \frac{GS_{G||} - S_{G\perp}}{(1 - 3l_2)GS_{G||} + (2 - 3l_1)S_{G\perp}}$$

where $S_{G||}$ and $S_{G\perp}$ are the average donor fluorescence count rates detected within a burst, with a polarization parallel and perpendicular to that of the excitation light, respectively. Correction factors, $l_1 = 0.0308$ and $l_2 = 0.0368$ were applied in order to take into account mixing of polarizations by the microscope objective lenses.[52] $G = 0.964$ compensates for slightly different detection efficiencies in the two detection channels.

The signal registered in the green detection range was described by scattered light and fluorescence decay with a single component. Thereby the fluorescence lifetimes were calculated using a maximum likelihood estimator as described in refs 28 and 53, which has been shown to be statistically most efficient. The standard deviations of the obtained fluorescence lifetimes of ±15% are equivalent to the expected shot noise limit, if the registered photon numbers and the fractions of scattered light were considered.

(52) Koshioka, M.; Saski, K.; Masuhara, H. *Appl. Spectrosc.* **1995**, *49*, 224−8.
(53) Brand, L.; Eggeling, C.; Zander, C.; Drexhage, K. H.; Seidel, C. A. M. *J. Phys. Chem. A* **1997**, *101*, 4313-21.

The results of the MFD measurements were presented by plotting the frequency of occurrence of certain pairs of parameter values in two-dimensional histograms, where increasing frequency is coded by the gray-scale level. To use the full dynamic range of MFD and to show also minor populations of molecules, we present *Z*-magnified views with a single peak level above a specified level indicated in yellow (the specific level is given in the captions of Figures 4−9). Raw data from several measurement sessions were combined into a single complete measurement set for further analysis.

Sample. For our studies, a deoxyoligonucleotide N1 containing 27 nucleobases in a random sequence was synthesized using the standard phosphoamidite technology: 5′-d(TTG AAA ACG AGA GAG ACA TAA ACG ATC). The 5′-d of this deoxyoligonucleotide was labeled via a C6 amino link either with the *N*-hydroxysuccinimide ester of Cy3 (Amersham Pharmacia Biotech) or Rh6G (Molecular Probes, Eugene, OR) as donor dye. Using a postlabeling reaction, the acceptor dyes BP650/665 (Molecular Probes) or Cy5 (Amersham Pharmacia Biotech) were covalently attached to an 5′-C6-amino-2′d**T** at a defined position (with 9, 13, or 17 base pairs (bp) between the linked dyes) on the complementary deoxyoligonucleotide N2: 5′-d(GAT CGT TTA **T**GT C**T**C TC**T** CGT TTT CAA). The oligonucleotides were double HPLC-purified and additionally PAGE-purified. All synthesis and purification was done by IBA GmbH (Göttingen, Germany). Hybridization was performed in a buffer containing 180 mM NaCl, 12 mM sodium citrate, and 25 μM $MgCl_2$ (pH 7.5). To obtain the FRET-active molecules, complementary donor- and acceptor-labeled strands (N1 and N2) were mixed in a 1:1 ratio and slowly cooled from 95 to 20 °C. In all experiments, we used a sodium phosphate buffer with 180 mM NaCl, 10 mM NaH_2PO_4/Na_2HPO_4, and 400 μM sodium ascorbate (pH 7.5). The following 16 molecules were analyzed: Cy3−Cy5 (9, 13, or 17 bp), Cy3−BP650/665 (9, 13, or 17 bp), Cy3−oligonucleotide, Rh6G−Cy5 (9, 13, or 17 bp), Rh6G−BP650/665 (9, 13, or 17 bp), Rh6G−oligonucleotide, and two fluorescent impurities of unknown structures **I1** and **I2**.

RESULTS AND DISCUSSION

Spectrofluorometric and Time-Correlated Single-Photon-Counting (TCSPC) Measurements. The labeled single-stranded oligonucleotides were investigated in a spectrofluorometer (Jobin-Yvon Fluorolog 3) with respect to the excitation and fluorescence spectra of the dyes attached to them. The fluorescence lifetimes were investigated by TCSPC, performed by the MFD instrumentation (Figure 2) at concentrations of ∼100 pM. In the TCSPC experiments, the donor (Rh6G, Cy3) and the acceptor (Cy5, BP650/665) dyes were excited by the argon ion (514 nm) and the krypton ion (647 nm) laser, respectively. The measured fluorescence lifetimes of the dyes and their maximum absorption and emission wavelengths are given in Table 1 and the spectra in Figure 3.

MFD Measurements and Strategy for Systematic Selection of Species. The different oligonucleotides were measured and analyzed burstwise with respect to all their fluorescence parameters on the single-molecule level by MFD, as described in Methods and Materials. For each type of dye-labeled oligonucleotide, the detected molecules were measured and sorted into a multidimensional cumulative histogram, based on the determined

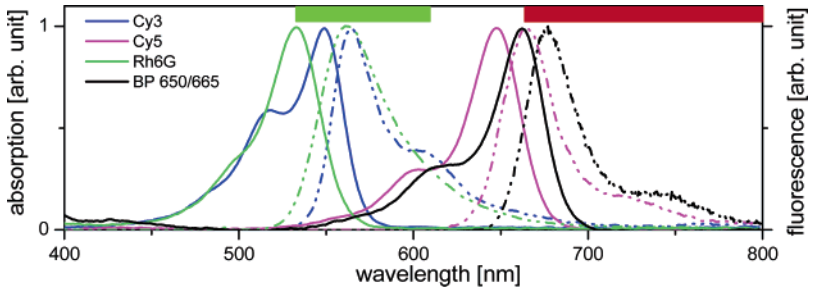

Figure 3. Absorption (solid line) and emission (dashed line) spectra of Rh6G, Cy3, Cy5, and BP650/665 dyes. Bars on the top indicate transmission range of the filters for single-molecule spectroscopy.

Table 1. Lifetimes and Abs/Flu Wavelengths

dye label at oligonucleotide	τ (ns)[a]	λ_{Abs}/λ_F (nm)	spectral detection (nm)
Rh6G	4.0	534/564	535−605
Cy3	0.8[b]	550/565	535−605
BP650/665	4.3	663/678	665−800
Cy5	1.0	649/665	665−800

[a] Fluorescence parameters of the free dyes are given. [b] The experimentally measured fluorescence lifetimes using the latest avalanche photodiodes (APDs) of the Series "AQR" are increased due to a count rate-dependent jitter by ~10%. The listed τ values are corrected for this effect, but the τ-histograms shown in Figures 4−9 were not corrected for this effect. [c] Surprisingly the Cy3-labeled DNA samples contain large amounts of a fluorescent species with a fluorescence lifetime of 1.5 ns. It is classified as an impurity because it is not quenched by FRET as the species with $\tau = 0.8$ ns. It is assumed that the species with $\tau = 1.5$ ns are not attached to the oligonucleotide.

fluorescence parameters. The fluorescence bursts from all the different oligonucleotides were then arranged into one joint histogram.

In Figure 4, a two-dimensional projection of this histogram is shown, where the fluorescence bursts from all the measured compounds have been arranged with respect to their count rates in the green and red emission ranges, S_G and S_R.

Even though recording takes place at the single-molecule level, which should in principle provide the possibility to reveal sub-

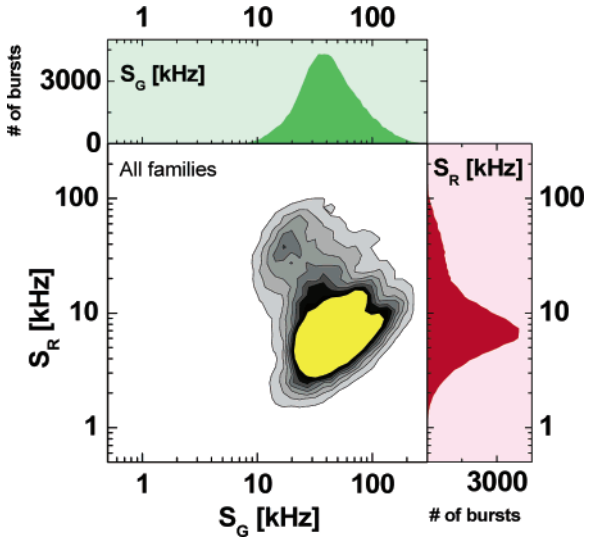

Figure 4. Cumulative histogram of all compounds. Green signal, S_G, is plotted versus red signal, S_R. Frequency is coded by gray scale. To visualize small subpopulations, a Z-magnified view is presented, where all frequencies above 0.2 of the maximum are displayed in yellow as a single level. Total number of molecules $A_{total} = 92\,694$.

populations within an ensemble of molecules, it is clear that at least a characterization based on only S_G and S_R is in this case not sufficient to identify the different subpopulations. From this histogram, it is also not possible to extract from how many different compounds the data have been recorded. To demonstrate the full dynamic range of MFD, we present Z-magnified views with the last level indicated in uniform yellow (see Figures 4 and 6−9), which focuses the dynamic range of the gray scale to the border zones between the different populations.

To better distinguish between the species, we follow a procedure where the subpopulations are identified based on a set of parameters and where a systematic selection of subpopulations within different ranges of the detected parameters will be performed in subsequent steps. The general rationale for this multistep selection procedure is that the different compounds differ distinctly from each other with respect to at least one measurable parameter or that combinations of parameter values can be used to enhance differences between compounds that do not differ distinctly with respect to a single parameter. In this study, differences between the compounds in rotational mobilities (characterized by different anisotropies, r_G, for free and labeled fluorophores), FRET efficiencies (different dye-to-dye distances reflected in S_G, S_R, τ_G, and τ_R), and differences due to the identities of the donor and acceptor dye labels (S_G, S_R, τ_G, τ_R), or differences in combinations of the parameters are exploited within the frame of this general strategy for systematic selection. As shown in Table 1, donor dyes (Rh6G, Cy3) and acceptor dyes (BP650/665, Cy5) with long and short fluorescence lifetimes, respectively, have been chosen for the green and the red spectral windows. Based on their characteristic fluorescence lifetimes, four distinct oligonucleotide families (Rh6G−BP650/665, Rh6G−Cy5, Cy3−BP650/665, Cy3−Cy5) with different DA distances have been studied.

Graphical MFD Analysis. Step 1: Selection in the r_G versus τ_G Parameter Plane. As a first step, the multidimensional cumulative histogram was projected onto a two-dimensional histogram, with the bursts sorted with respect to the measured fluorescence anisotropy, r_G, and the fluorescence lifetimes, τ_G, of their donor fluorophores ($r_G-\tau_G$ histogram). Given its formal expression (see Methods and Materials), it is clear that r_G will increase with shorter fluorescence lifetimes. This is illustrated in the $r_G-\tau_G$ histogram of Figure 5. For comparison, lines corresponding to mean rotational correlation times ρ of 0.2, 2.0, and 7.0 ns are included, which are calculated from the Perrin equation, $r_G = r_0/(1 + \tau/\rho)$, using a fundamental anisotropy $r_0 = 0.375$.

In view of the broad range of fluorescence lifetimes covered, the observed anisotropy, broadened by shot noise, is also very

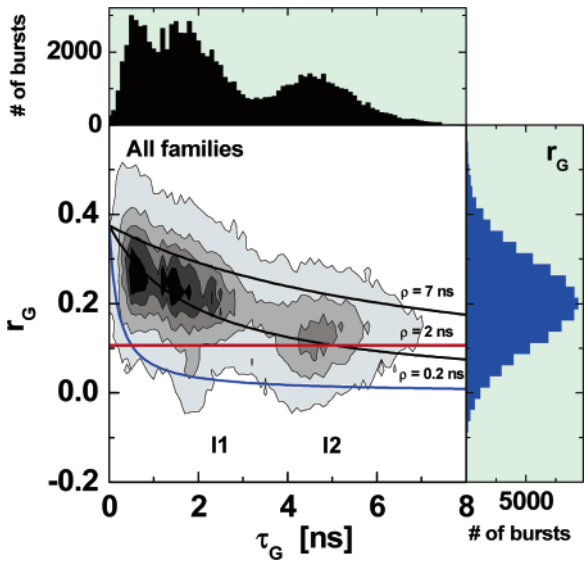

Figure 5. Green lifetime, τ_G, plotted versus anisotropy, r_G, together with an overlaid Perrin equation computed for three rotational correlation times (0.2, 2, 7 ns). Molecules under the red line were excluded from further analysis. Total number of molecules A_{total} = 92 628.

wide. The anisotropies of the major subpopulations consisting of the donor dyes linked to the oligonucleotides are well described by a mean rotational correlation time of $\rho = 2.0$ ns. However, there are two subpopulations, marked as **I1** and **I2**, with fluorescence lifetimes $\tau_G = 1.8$ and 4 ns, which have very small mean rotational correlation times of ~0.2 ns (blue line). The small rotational correlation time is evidence for the existence of free impurities, which emit in the green spectral range and are not subject to FRET-mediated quenching of their fluorescence.

Based on this, a limit was defined such that only labeled oligonucleotides with $r_G > 0.107$ (red line) were selected for further analysis. The highly rotationally mobile fluorescent impurities **I1** and **I2** could thus be sorted out with high certainty. Their rotational correlation times were short enough not to be attributed to the rotation of donor-labeled oligonucleotides, but rather to free dyes.

Generation of Characteristic Patterns. In the present qualitative analysis, we aim at detection of the presence of an individual subpopulation. For this, we generated characteristic patterns of different individual species of donor−acceptor-labeled oligonucleotides. These characteristic patterns guide the definition of the areas in various parameter planes as belonging to a particular species. We generated specific patterns for all two-dimensional parameter planes by analyzing the individual subpopulations separately. In addition to free dye molecules, fractions of molecules that are not labeled at all or that are not completely labeled remain in practically all labeling procedures. To show the features distinguishing fully labeled molecules from incompletely labeled ones, we refer to the cumulative histograms of the individual oligonucleotide species from which free dye molecules have been sorted out by use of the criterion of step 1. Panels A−C in Figure 6 show three typical two-dimensional histograms, with oligonucleotides labeled with Rh6G (donor) and Cy5 (acceptor) with different distances between the labeling sites. The measurements can be evaluated by depicting the ratios of their signal intensities detected in the green (S_G) divided by those in the red (S_R) wavelength range, S_G/S_R, along the ordinate, and their donor fluorescence lifetimes, τ_G, along the abscissa.

It can be seen from these $S_G/S_R-\tau_G$ histograms that the fractions of oligonucleotides that are not labeled with acceptor dyes can be clearly distinguished from those labeled with both dyes. For oligonucleotides that were labeled with both donor and acceptor dyes (DA-oligonucleotides), FRET-mediated quenching of the donor fluorescence and a concomitant FRET-sensitized fluorescence of the acceptor dye will occur when the donor dye is excited by laser light at 514 nm. With higher FRET efficiencies, increasingly lower ratios of S_G/S_R and shorter τ_G appear. DA-oligonucleotides accumulate in distinct fractions with lower S_G/S_R and shorter τ_G, well separable from oligonucleotides, having just a donor fluorophore labeled to them (D-oligonucleotides). Acceptor-labeled oligonucleotides (A-oligonucleotides) are practically not excited at an excitation wavelength of 514 nm and are thus not detectable in the current experiment.

Further, for DA-oligonucleotides labeled with the same donor−acceptor dye pair, the positions of the population of DA-oligo-

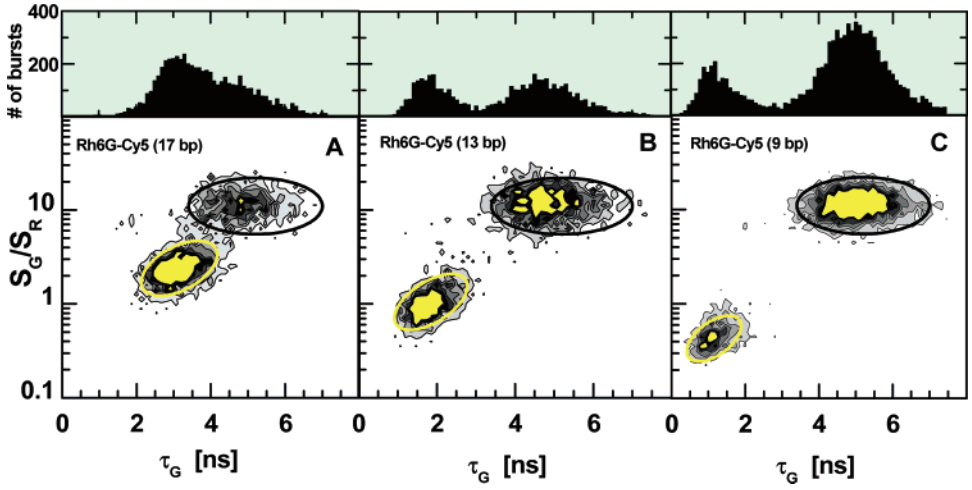

Figure 6. Green lifetime, τ_G, plotted versus fluorescence count rate ratio, S_G/S_R, for individual populations of oligos labeled with Rh6G (donor) and Cy5 (acceptor) with different distances between the labeling sites [17 (A), 13 (B), and 9 base pairs (C)]. Black ellipses show patterns for donor-only-labeled oligos, yellow ellipses for oligos labeled with both donor and acceptor. Z-Magnified view, where all frequencies above 0.4 of the maximum are displayed in yellow as a single level. Total number of molecules: (A) A_{total} = 5856, (B) A_{total} = 4768, and (C) A_{total} = 9804.

nucleotides in the histograms differ distinctly depending on the dye-to-dye distance (distances between the two dyes in Figure 6A, B, and C of 17, 13, and 9 base pairs, respectively). The differences of the individual histograms should thus provide patterns from which it is possible to separate D-oligonucleotides from DA-oligonucleotides and to distinguish between the DA-oligonucleotides in the joint histograms via their specific FRET efficiencies defined by the distances between the labeling sites of their donor and acceptor fluorophores.

For this graphical analysis, it is sufficient to approximate the characteristic patterns by variable ellipses with characteristic colors for the four families of DA oligonucleotides [Rh6G−BP650/665 (magenta), Rh6G−Cy5 (yellow), Cy3−BP650/665 (orange), Cy3−Cy5 (olive)] and for the D-oligonucleotides of Rh6G and Cy3 (black). The size of these elliptic patterns in the contour plots is defined by the area in which the frequency is 20% of the maximum value or higher. For histograms containing a mixture of species, such elliptic patterns guide the definition of areas within which the different individual species are confined.

Step 2: Selection in the S_G/S_R versus τ_G Parameter Plane. After separation of free dyes from the mixture in step 1 above, differences in the macroscopic fluorescence properties of the fluorophores (Table 1) were exploited in sequential steps to separate the subpopulations of the remaining oligonucleotides. First, the oligonucleotides were separated into two main groups based on the differences in the intrinsic fluorescence lifetimes of their donor fluorophores [Rh6G (4 ns) and Cy3 (0.8 ns)]. This was performed by displaying the detected oligonucleotides in a S_G/S_R versus τ_G histogram. As mentioned above, higher efficiencies of FRET generate increasingly lower ratios of S_G/S_R and shorter τ_G. However, irrespective of the degree of quenching of the donor dye due to FRET, oligonucleotides labeled with Rh6G can be separated from those labeled with Cy3. This separation is possible by virtue of the distinct differences in the measured τ_G between the molecules, which follows from the prominent difference in the intrinsic lifetimes of the two donor fluorophores. The separation of Cy3-labeled oligonucleotides from their Rh6G-labeled counterparts is indicated by a red line in Figure 7.

Step 3: Selection in the S_G/S_R versus τ_R Parameter Plane. Following the separation of oligonucleotides based on the intrinsic donor fluorescence lifetimes (Figure 7), oligonucleotides labeled with either Cy3 or Rh6G as donors were separated with respect to what acceptor dye was attached to them. This separation step is possible due to a large difference in the intrinsic lifetimes of the two acceptor dyes Cy5 (1.0 ns) and BP650/665 (4.3 ns) (see Table 1). In Figure 8A and B, the histograms selected based on the different intrinsic donor lifetimes of the Cy3- and Rh6G-labeled oligonucleotides according to step 2 are each projected in a plane defined by the ratio S_G/S_R and the measured decay time in the red detection channels, τ_R. Here, a clear distinction can be seen for those oligonucleotides labeled with BP650/665 (4.3 ns) as the acceptor dye from those with Cy5 (1.0 ns). Disregarding the effect of the duration of the laser excitation pulses (which is subject to deconvolution), the measured acceptor fluorescence intensity is a convolution product of the donor and acceptor fluorescence intensity decay following a Dirac pulse, i.e.,

$$F_A(t) \propto \int_0^t \exp(-t'/\tau_D)\, \exp(-(t-t')/\tau_A)\, dt'$$

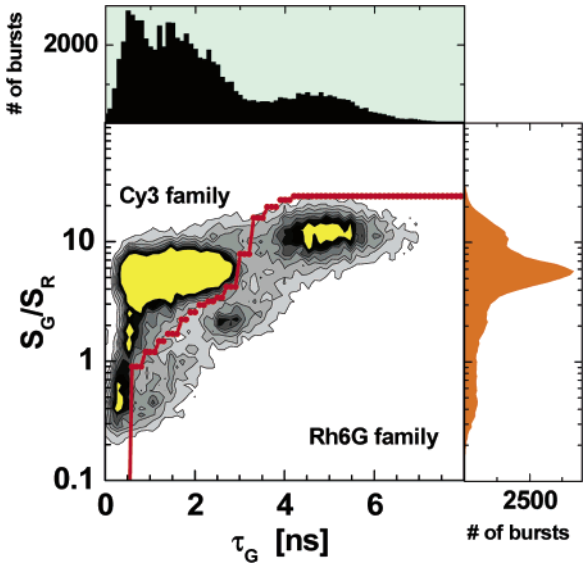

Figure 7. Green lifetime, τ_G, plotted versus fluorescence signal ratio, S_G/S_R, for all measured compounds. Red line shows separation between populations of oligos labeled with Rh6G and Cy3 dyes. *Z*-Magnified view, where all frequencies above 0.2 of the maximum are displayed in yellow as a single level. Total number of molecules $A_{total} = 74\,758$.

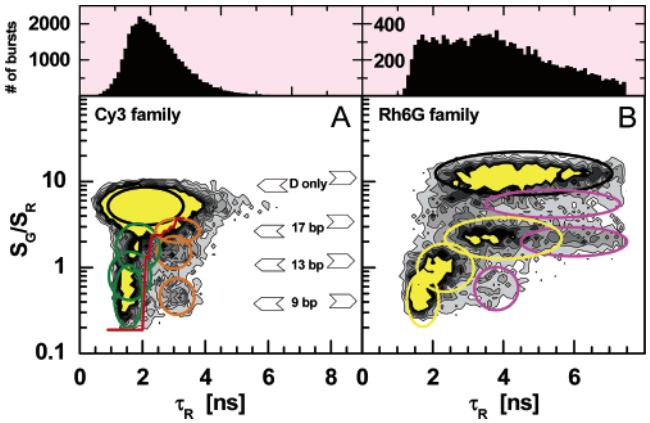

Figure 8. Red decay time, τ_R, plotted versus fluorescence signal ratio, S_G/S_R, for oligos, labeled with Cy3 (A) and Rh6G (B) donor dyes. Green ellipses indicate patterns for Cy3−Cy5 populations; orange, Cy3−BP650/665 populations; yellow, Rh6G−Cy5 populations; pink, Rh6G−BP650/665 populations; black, donor-only-labeled oligos. Red line shows separation between Cy3−BP650/665 and Cy3−Cy5 populations. *Z*-Magnified view, where all frequencies above 0.075 of the maximum are displayed in yellow as a single level. Total number of molecules: (A) $A_{total} = 37\,275$; (B) $A_{total} = 14\,398$.

for monoexponential fluorescence decays, where τ_D and τ_A are the lifetimes of the donor and acceptor dyes, respectively. In contrast to the determination of the donor fluorescence lifetime from the complete fluorescence response curve, it is at present sufficient that only the decaying part of the complex fluorescence response of the acceptor is judged by a single-exponential decay without deconvolution having a characteristic signal decay time, τ_R.[54] Short D−A distances generate short donor lifetimes and, thus, also a shorter time over which excitation by energy transfer to the acceptor dye can take place. Consequently, for oligonucle-

(54) Eggeling, C.; Schaffer, J.; Seidel, C. A. M.; Korte, J.; Brehm, G.; Schneider, S.; Schrof, W. *J. Phys. Chem. A* **2001**, *105*, 3673−9.

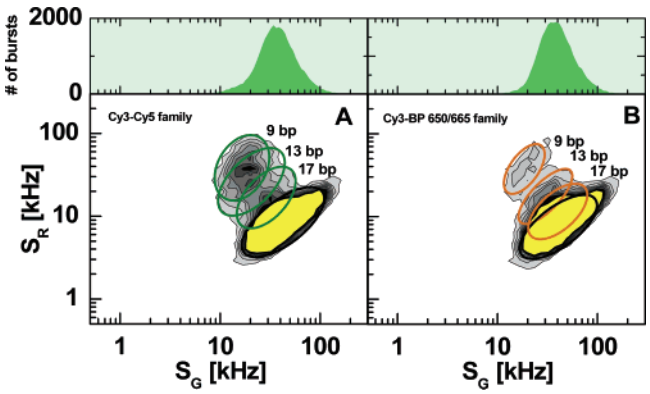

Figure 9. Green fluorescence signal, S_G, plotted versus red signal, S_R, for Cy3−Cy5 populations (A) and Cy3−BP650/665 populations (B). Green ellipses indicate patterns for Cy3−Cy5 populations and orange, Cy3−BP650/665 populations. *Z*-Magnified view, where all frequencies above 0.06 of the maximum are displayed in yellow as a single level. Total number of molecules: (A) A_{total} = 29 035; (B) A_{total} = 29 306.

otides with shorter D−A distances, typically a shorter time interval passes after the excitation pulse until the acceptor dye is excited and subsequently fluoresces, although the intrinsic fluorescence lifetime of the acceptor remains the same. This effect can most clearly be seen for those oligonucleotides having Rh6G as a donor dye (Figure 8B), since the absolute changes in fluorescence lifetime upon FRET are larger with Rh6G than with Cy3, due to the shorter intrinsic fluorescence lifetime of Cy3. Likewise, with a relatively short intrinsic fluorescence lifetime of the acceptor dye, the relative effect of convolution by the donor fluorescence decay is more prominent and distinct, as can be seen for the Rh6G−Cy5-labeled oligonucleotides in Figure 8B. Considering the specific patterns, all DA combinations of the Rh6G family can easily be distinguished, which allows for the simultaneous detection of seven oligonucleotides: Rh6G−Cy5 (9, 13, and 17 bp), Rh6G−BP650/665 (9, 13, and 17 bp), and Rh6G−oligonucleotide (D-oligonucleotide).

Step 4: Selection in the S_R versus S_G Parameter Plane. As indicated in Figure 8A, a separation relying on D−A distances can be difficult when based solely on the S_G/S_R versus τ_R projection. Considering the Cy3 family, a substantial overlap of the oligonucleotides with different D−A distances can be noticed. For the case of overlap between populations of different species, additional projections can be used to increase the specificity of the separation of the different species from each other. In Figure 9A and B, the projections of the Cy3−Cy5 and Cy3−BP650/665 populations, both included in Figure 8A, are after separation (indicated by the red line in Figure 7A) plotted along S_G and S_R. In this case, this projection provides an additional criterion, enabling an improved separation of the oligonucleotides labeled with Cy3 as a donor and either Cy5 (olive) or B650 (orange) as an acceptor, based on their D−A distances. In this way, the remaining seven oligonucleotides are identified: Cy3−Cy5 (9, 13, and 17 bp), Cy3−BP650/665 (9, 13, and 17 bp), and Cy3−oligonucleotide (D-oligonucleotide).

Quantitative Identification of the Detected Molecules Using a Pattern Recognition Algorithm of Fluorescence Parameter Sets. The visual identification procedure, performed in sequential steps, as outlined above, forms the principal basis of an automated identification procedure that can determine the relative fractions of each species. The accuracy of such an approach would ideally be limited only by the ability to identify properly single-molecule bursts for each species in the data. In the following section, we present a procedure, which exploits the multidimensionality of the measured data and which lends itself well for fast, stable, and reliable species identification of the detected species.

Procedure. The core of the procedure is the estimation of the probability, $\Psi(m|\{v_{k\,\mathrm{exp}}\}_\alpha)$. It states the probability that a detected molecule α, possessing a set of K independent experimentally determined fluorescence parameter values, $\{v_{1\,\mathrm{exp}}\ldots v_{K\,\mathrm{exp}}\}$, belongs to species m out of N different species with the index $n = 1$ to N. In the K-dimensional set $\{v_{k\,\mathrm{exp}}\}_\alpha$, each of the K values corresponds to a unique fluorescence parameter, e.g., lifetime, anisotropy, etc.

As an intermediate step in determining Ψ, it is first necessary to determine a set of $K \times N$ probability distributions, $\{P_{k,m}(v_k)$, $k = 1$ to K, $n = 1$ to $N\}$, which describes the probability that a molecule known to be of species m has a value for the kth fluorescence parameter given by v_k (or more precisely, has a value that falls into a bin centered around v_k). In this report, the $P_{k,m}(v_k)$ distributions are generated from histograms obtained from the same MFD data used to determine the characteristic patterns above (for example, see 1D projections in Figure 6), whereby the DA-oligonucleotides were first separated from the D-oligonucleotides and impurity species. The relevant histograms are normalized such, that for each species m and parameter k, $\sum_v P_{k,m}(v_k) = 1$ (remembering that the values for v are restricted to a finite set of bin centers). The probability $\Psi(m|\{v_{k\,\mathrm{exp}}\}_\alpha)$ is calculated from the set of distributions $\{P_{k,m}(v_{k\,\mathrm{exp}})\}$ as follows:

$$\Psi(m|\{v_{k\mathrm{exp}}\}_\alpha) = \frac{\prod_{k=1}^{K} P_{k,m}(v_{k\mathrm{exp}})}{\sum_{n=1}^{N}\prod_{k=1}^{K} P_{k,n}(v_{k\mathrm{exp}})} \qquad (1)$$

This ratio expresses the probability of a molecule from species m displaying the particular set of parameters $\{v_{k\,\mathrm{exp}}\}$ (calculated in the numerator) relative to the probability that the set $\{v_{k\,\mathrm{exp}}\}$ will be displayed by *any* of the N species (denominator). In the absence of conformational dynamics (which is here the case), shot noise limited distributions for the individual parameters are measured. For each individual molecule α, a set of N probabilities, $\{\Psi(1|\{v_{k\,\mathrm{exp}}\}_\alpha)\ldots \Psi(N|\{v_{k\,\mathrm{exp}}\}_\alpha)\}$, is calculated by eq 1 using the set of experimentally determined fluorescence parameters for that molecule, $\{v_{1\,\mathrm{exp}}\ldots v_{K\,\mathrm{exp}}\}_\alpha$. Since a molecule must belong to one of the N species, $\sum_n \Psi(n|\{v_{k\,\mathrm{exp}}\}_\alpha) = 1$ for each molecule α. The molecule is then classified as belonging to that species m for which $\Psi(m|\{v_{k\,\mathrm{exp}}\}_\alpha)$ is the highest.

For each of the $n_d = 1, \cdot, m, \cdot, N$ DA-oligonucleotide data sets, a histogram was generated that counted how many molecules in that data set were assigned to each of the N species. By normalizing the sum of each of these histograms to unity, we obtain the classification probability $\Omega(m|n_d)$ stating the probability that a molecule of the data set n_d (i.e., originating from species $n = m$) is classified as species m. By restricting analysis to a single data set n_d, the classification probability $\Omega(m|n_d)$ can be defined

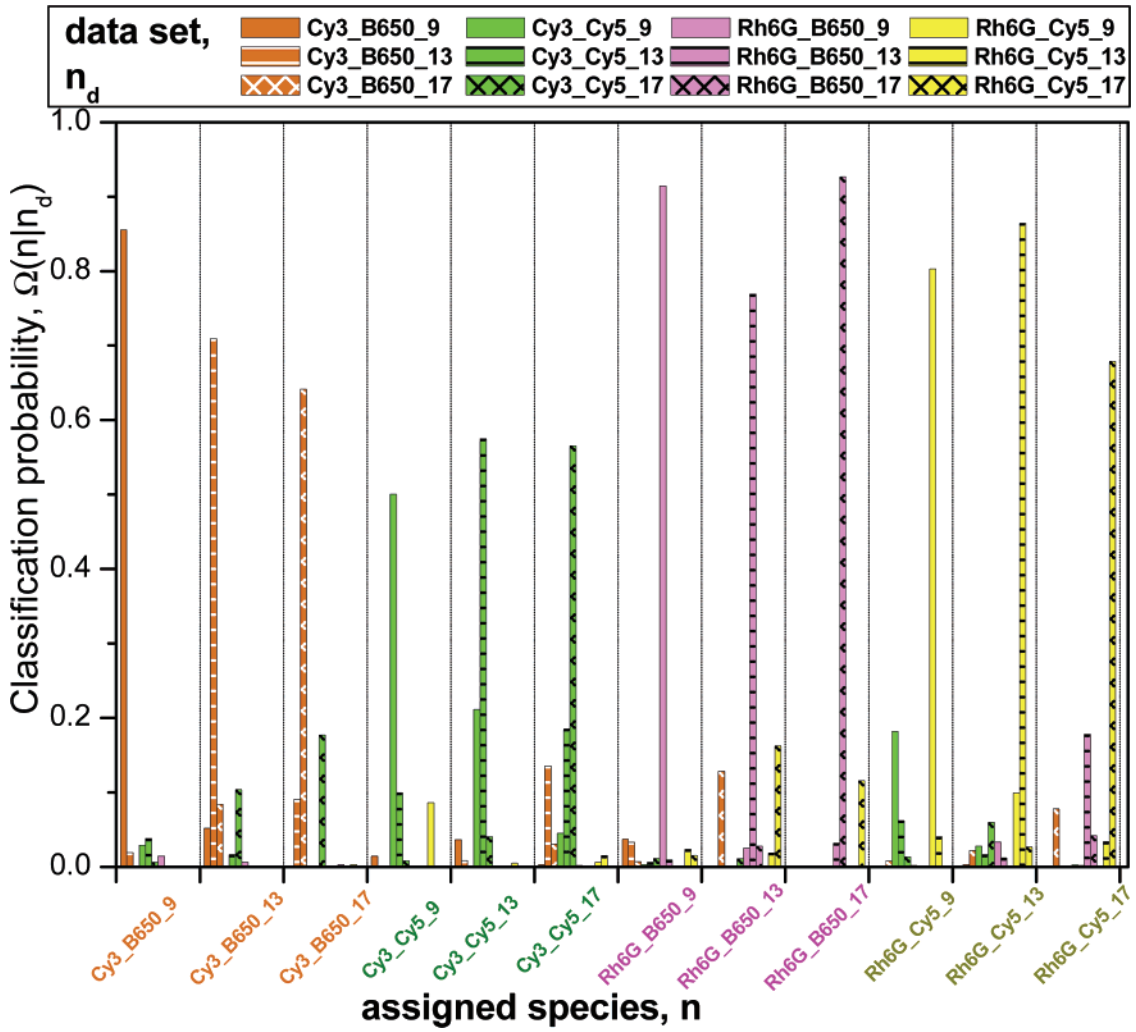

Figure 10. Classification probability, $\Omega(n|n_d)$, that an DA-oligonucleotide molecule of the data set n_d is classified as a species n. Calculations were based on the procedure, described in the text, using four fluorescence dimensions: green lifetime, τ_G, green anisotropy, r_G, red decay time, τ_R, and signal intensity ratio, S_G/S_R. (For detailed values see Supporting Information, Table 1).

as $\Omega(m|n_d) = a_m/\sum_{n=1}^{N} a_n$, where $\{a_{n=1}; \ldots a_N\}$ are the numbers of molecules from data set n_d assigned to species m.

Classification of Single Molecules. In Figure 10, the classification probabilities obtained for each of the $N = 12$ DA-oligonucleotide species are shown. If $\Omega(n=n_d|n_d) = 1$, then we have no errors in classification. However, not all of the molecules of an individual data set n_d are indeed classified as belonging to species m, i.e., $\Omega(m=n_d|n_d) < 1$. This error is due to overlap of the one-dimensional probability distributions, $P_{k,n}$, between the different species. However, by increasing the number of dimensions, K, the classification confidence can be improved. The classification probabilities shown in Figure 10 use four fluorescence dimensions: green lifetime, τ_G, green anisotropy, r_G, red signal decay time, τ_R, and signal intensity ratio, S_G/S_R. Analyzing $\Omega(m|n_d)$ in Figure 10, the likelihood that a molecule of data set is properly assigned to a species (i.e., the correct classification, $\Omega(n=n_d|n_d)$ always dominates over the other classification probabilities $\Omega(n\neq n_d|n_d)$ and exceeds 0.5 for all data sets with n_d.

Analysis of an Ensemble of Molecules. In general, having a definition for the classification probabilities $\Omega(n=1|n_d)$ to $\Omega(n=N|n_d)$ for a given data set n_d with a separate species n allows us to generalize the definition for a_n to include all of the analyzed data sets. The frequency histogram for the classification of corresponding single molecules can be predicted with the frequency of the assigned molecules described by $\vec{a}=\{a_{n=1}; \ldots a_N\}$, where the total number of detected molecules A_{total} is given by $\sum_{n=1}^{N} a_n$. Due to shot noise broadened fluorescence parameters, the patterns overlap, which has the consequence that the recognition of single molecules cannot be perfect (Figure 10). Because of these statistical classification problems in the first step, which can be described by error probabilities, we now analyze all molecules together in a second step and improve the classification on the basis of the knowledge of the mean classification error. In this way, a much better result for the number of detected molecules of the ensemble c_n is obtained than at the single-molecule level.

The actually calculated numbers for each species in this mixture $\vec{c} = \{c_{n=1}; \ldots c_N\}$ are obtained by solving the following system of linear equations with the classification probabilities $\Omega(n|n_d)$ and the experimentally obtained numbers of assigned molecules $\vec{a}$.

$$\begin{bmatrix} \Omega(n=1|n_d=1) & \ldots & \Omega(n=1|n_d=N) \\ \ldots & \ldots & \ldots \\ \Omega(n=N|n_d=1) & \ldots & \Omega(n=N|n_d=N) \end{bmatrix} \begin{bmatrix} c_{n=1} \\ \ldots \\ c_N \end{bmatrix} = \begin{bmatrix} a_{n=1} \\ \ldots \\ a_N \end{bmatrix} \quad (2)$$

Finally, we want to give an example for the substantial improvement by using eq 2 to reliably calculate ensemble numbers of the analyzed molecules. Therefore, we analyzed an independently measured test sample with a total number $A_{total} = 4143$ molecules, which contained three of the 12 DA-oligonucleotides. For the discussion of the results we normalize the molecule numbers $\vec{a}$ and $\vec{c}$ by A_{total} defining relative fractions of experimentally assigned molecules $\vec{a}_r$ and relative species fractions calculated for the ensemble $\vec{c}_r$, respectively. The actual sample composition $\vec{c}_{r\text{sample}} = [a_3(\text{Cy3}-\text{BP650/665 (17 bp)}) = 0.33;\ a_4(\text{Cy3}-\text{Cy5 (9bp)}) = 0.35;\ a_{12}(\text{Rh6G}-\text{Cy5 (17bp)}) = 0.32\}$ is very well recovered by our single-molecule counting analysis $\vec{c}_{r\text{analysis}} = \{a_3 = 0.32;\ a_4 = 0{,}35;\ a_6 = 0.01;\ a_8 = -0.01;\ a_{11} = -0.01;\ a_{12} = 0.34\}$. Please note that in the following analysis we list only coefficients for the relative species fractions, which are not equal to zero. The negative fractions of $\vec{c}_{r\text{analysis}}$ are solely the result of the mathematical solution and have of course no chemical meaning. In contrast to $\vec{c}_{r\text{analysis}}$, please note that the fractions obtained in the first step by the single-molecule classification $\vec{a}_r = \{0.01;\ 0.03;\ \mathbf{0.21},\ \mathbf{0.17};\ 0.08;\ 0.03;\ 0.01;\ 0.09;\ 0.04;\ 0.06;\ 0.02;\ \mathbf{025}\}$ (eq 1) indicate a level of misclassification of up to 10%, which results in false positive and false negative classified molecules on the single-molecule level. This result is equivalent to the case of the overlap of ellipses for different species in the graphical analysis (e.g., Figure 8). However, using eq 2, $\vec{a}_r$ is converted with high accuracy to $\vec{c}_{r\text{analysis}}$ in the second step of the single-molecule counting analysis. The comparison of $\vec{c}_{r\text{sample}}$ with $\vec{c}_{r\text{analysis}}$ indicates a very low statistical noise level of only 1−2%.

CONCLUSIONS AND OUTLOOK

The purpose of this study was to demonstrate the feasibility and huge dynamic range of single-molecule measurements for identifying a larger number of different species, based on multiparameter detection and a multidimensional single-molecule counting analysis. By MFD, free dye molecules and incompletely labeled species could be excluded from the analysis, and in all 16 different compounds (four families of DA-oligonucleotides, Cy3−Cy5, Cy3−BP650/665, Rh6G−Cy5, and Rh6G−BP650/665, each with 9-, 13-, and 17-bp separation between the dyes, D-oligonucleotides labeled with Cy3 or Rh6G, impurities) were separated, based on the multidimensional information contained in the fluorescence of the analyzed molecules.

Following arrangement of the detected molecules into a multidimensional cumulative histogram, identification of the different species was first demonstrated by use of sequential two-dimensional projections of the histogram in different parameter planes, employing different "borderlines" between the populations projected in each plane. However, apart from this visual approach demonstrating the feasibility, we also show that a more quantitative identification procedure can be employed, based on an algorithm that offers fast and automated species identification with reasonable fidelity on a single-molecule level. This approach can also regenerate concentration profiles of species in a mixed sample with very high accuracy, and there is still a clear potential for implementing additional criteria and procedures. Apart from the "traditional" "chromophore" fluorescence parameters providing the criteria for the species identification in this study, fluctuation analysis of the fluorescence intensity (such as fluorescence correlation spectroscopy) or some other parameter of the detected fluorescence) can yield additional independent parameters. To such characteristic fluctuation parameters, within the category of parameters referred to as "environment/system parameters" above, belong triplet-state transitions,[55] trans−cis isomerization,[56] charge transfer,[57] or even transport properties.[54] In a recent study, the determination of up to 14 different parameters from one individual molecule was claimed, including transition rates to and from the triplet state and spatial parameters.[58] In MFD, fluctuation analysis of the detected fluorescence within selected bursts is possible. The use of MFD and additional parameters extracted from fluctuation analysis to identify different individual molecules will be reported elsewhere. Recently, Bowen et al. implemented neural networks for classifying fluorescence from individual molecules in a mixture. A neural network was constructed and trained to identify molecules based on fluorescence lifetime data and was shown to perform better than (nonideal more complex data) or equal to (ideal data) a maximum likelihood estimator (MLE) approach.[23] Here "ideal" means, that the data could be well described by the MLE by assuming, for example, a single-exponential decay. By use of identification procedures of this kind, applied in the multiparameter space, it is reasonable to believe that the number of species that can be separated from each other, and the specificity with which this can be done, can be increased even further. Apart from extensions in the extraction and processing of information from the emitted fluorescence, further operations can also be added on the excitation side, extending the degrees of freedom with respect to which single molecules can be analyzed. Recently, Kapanidis et al. reported the use of alternating direct laser excitation of both donor and acceptor fluorophores on single diffusing molecules. This enables both an interdye distance-dependent and distance-independent intensity ratio of the analyzed molecules, the latter based on labeling stoichiometry.[59] MFD lends itself well for the introduction of an alternating excitation scheme, with two or several laser lines. In general, given the flexible concept of MFD, in terms of possible schemes of excitation, the emission characterization, and the data processing steps, the potential for further improvements is large, and the modifications can be tailored to the specific species to be identified. The presented study thus only indicates a fundamental level at which single-molecule detection and identification can be performed using the MFD approach.

Finally, it is important to mention that an analysis of a sample, where the analyzed fluorescence bursts have been selected from the fluorescence intensity time trace via an intensity threshold criterion, can give only a qualitative picture on the sample composition, because the chance of bright molecules to be selected and to be subsequently analyzed is higher than that of less bright molecules. However, the multidimensional single-molecule counting analysis performed here is very well suited to check for the presence of even very small amounts of analyte molecules yielding a qualitative sample survey. Moreover, Fries et al.[49] have shown that the analysis of the obtained burst size distributions allows one to identify fluorescent sample molecules

(55) Widengren, J.; Mets, Ü.; Rigler, R. *J. Phys. Chem.* **1995**, *99*, 13368−79.
(56) Widengren, J.; Schwille, P. *J. Phys. Chem. A* **2000**, *104*, 6416−28.
(57) Widengren, J.; Dapprich, J.; Rigler, R. *Chem. Phys.* **1997**, *216*, 417−26.
(58) Prummer, M.; Sick, B.; Renn, A.; Wild, U. P. *Anal. Chem.* **2004**, *76*, 1633−40.
(59) Kapanidis, A. N.; Lee, N. K.; Laurence, T. A.; Doose, S.; Margeat, E.; Weiss, S. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 8936−41.

also in a quantitative manner. Thus, the introduced method of multidimensional single-molecule counting harbors the potential for a complete characterization and quantitative analysis of a highly diluted sample in homogeneous assays.

ACKNOWLEDGMENT

Financial support by BMBF Biofuture 0311865 is gratefully acknowledged. Enno Schweinberger is acknowledged for assisting with the design of oligonucleotides. Suren Felekyan, Stanislav Kalinin, and Wajih Al-Soufi are acknowledged for helpful theoretical discussions and for providing important analytical software. J.W. acknowledges support from the Swedish Board for Internationalization of Higher Education and Science (STINT), The Swedish Research Council (VR), and the Swedish Cancer Foundation.

SUPPORTING INFORMATION AVAILABLE

Additional information as noted in text. This material is available free of charge via the Internet at http://pubs.acs.org.

Received for review December 23, 2005. Accepted January 3, 2006.

AC0522759