Measuring number of fluorophores labelling cDNA strands, in solution, with Fluorescence Correlation Spectroscopy and photobleaching

Antoine Delon, Irène Wang, Emeline Lambert, Silva Lerbs-Mache, Régis Mache, Jacques Derouard, Vincent Motto-Ros, Rémi Galland

Measuring number of fluorophores labelling cDNA strands, in solution, with Fluorescence Correlation Spectroscopy and photobleaching

Antoine Delon[a]*, IrèneWang[a], Emeline Lambert[b], Silva Mache[b], Régis Mache[b], Jacques Derouard[a], Vincent Motto-Ros[a],[c], Rémi Galland[a]

[a] Univ. Grenoble I / CNRS, Laboratoire de Spectrométrie Physique UMR 5588, BP 87, 38402 Saint Martin d'Hères

[b] Univ. Grenoble I / CNRS, Laboratoire de Physiologie Cellulaire et Végétale UMR 5168, BP 53, 38041 Saint Martin d'Hères

[c]Present adress: Univ. Lyon I / CNRS, LaSIM UMR 5579, 69622 Villeurbanne Cedex

* corresponding author : adelon@ujf-grenoble.fr

Abstract

We present different approaches that aim at determining, in solution, the brightness and the number of Alexa Fluor 647 molecules labelling the C bases of two sequences of cDNA, corresponding to two transcripts of different sizes, a short and a long transcript (123 and 306 base long, with 45 and 74 dCTP residues, respectively). In each case, the Alexa labeled bases have been incorporated during reverse-transcription. Two kinds of experiments have been performed and combined: photobleaching and Fluorescence Correlation Spectroscopy (together with the factorial cumulant analysis method). As a result, we show that the photobleaching cross-section of labelled cDNA strands is about half that of free Alexa in aqueous solution, while their brightness is about twice. The factorial cumulant analysis put into evidence the fact that the brightness of cDNA strands varies from molecule to molecule, due to the statistical distribution of the number of Alexa fluorophores labelling cDNA. Our measurements are consistent with a poissonian distribution of the number of fluorescent labels per cDNA, with a mean value of about 2 for the long cDNA strands. This indicates that only a few percent of the dCTP incorporated during the reverse transcription hold a fluorescent label, which is consistent with the fact that these Alexa Fluor 647 labels do not interact with each other. Finally, the analysis of experiments also reveals that the photobleaching decay rate of the excited state of Alexa Fluor 647 decreases by about 30% when incorporated into cDNA, while its non radiative decay rate is increased such that the brightness of individual Alexa labels is decreased by 25%.

Keywords: DNA, Fluorescence Correlation Spectroscopy, labelling, photophysics, single molecule studies

1. Introduction

The multiple labelling of biological molecules, such as proteins or DNA strands, with fluorophores, is a key point for molecular imaging in microscopy, the investigation of biomolecular processes,[1, 2] biochips, DNA sequencing,[3, 4, 5, 6, 7] *etc*… The knowledge of the number of fluorophores incorporated into the molecules is an important issue since it determines the brightness of the molecules to be detected and affects the detection limit. This is especially true for single molecule measurements,[8] that have, for instance, revealed aggregation that would have not been detected otherwise.[9] When the goal is to measure a single molecule, it is important to be able to distinguish aggregates of various orders, which implies a proper understanding of multiple labelling. Especially demanding are experiments that aim at quantifying aggregation of molecular systems.[10, 11] One possible method to measure the number of dyes present in an aggregate is to count the number of steps observed in the time trace of the fluorescence of single molecules before they are completely photobleached.[1, 12] More sophisticated methods have also been proposed, that rely on a statistical analysis of the fluorescence trajectories during photobleaching or on the deconvolution of the intensity distribution of single molecules.[13, 14] However, all of them require the molecules to be immobilized on a substrate and observed using a high detection efficiency and specificity and to analyse a large number of cases to get the whole statistical distribution of labelling and brightness. This is not compatible with any biological systems, so that bulk experiments that would nevertheless permit to access information at a single molecule level would be of high interest. This is in fact the case of Fluorescence Fluctuation Spectroscopy and related methods.[15] Although this is not, strictly speaking, a single molecule technique, it exploits the statistical properties of the fluorescence fluctuations from a limited amount of molecules present at a time in the observation volume.

Here we consider the case of cDNA strands in solution. This considerably simplifies the experiment. Although the measurements are primary averaged over this ensemble, we will show that it is still possible to extract further information concerning the nature of the statistical distribution of labelling.

Our method relies heavily on the measurement of the effective brightness of the cDNA molecules in solution, defined as the ratio of the total fluorescence signal to the effective mean number of molecules in the observation volume and on the analysis of its behavior while the cDNA strands are progressively photobleached. This has been carried out using two main tools: i) Fluorescence Correlation Spectroscopy (FCS) to measure the effective mean number of molecules in the observation volume,[16] together with their brightness and ii) photobleaching of the molecules confined in small volumes (microwells imprinted in polydimethylsiloxane). Additionally, Fluorescence Fluctuation Spectroscopy has been performed, using statistical cumulant analysis, to confirm the statistical nature of the number of fluorophores labelling the cDNA strands.[17] This latter approach has the other advantage of offering a powerful theoretical framework to interpret the brightness measurements.

The experiments have been performed on single cDNA strand having Alexa-Fluor-647-dCTP incorporated during reverse transcription. The data are then analysed and interpreted in terms of multiple labelling, using three models of increasing complexity, that combine FCS and photobleaching experiments: *Model 0*, based on the photobleaching cross-section and brightness of the cDNA strands, assumes that all the strands have all the same number of Alexa labels; *Model 1*, incorporates the statistical nature of the multiple labelling; *Model 2*, independent of any hypothesis concerning the photophysical parameters, takes profit of the variations of the molecular brightness and number of molecules when sequentially photobleaching the cDNA sample. This progressive approach makes it possible to understand the limitations of the simpler ways of thinking and the necessity of a more sophisticated ones.

In addition, our analysis gives some indications concerning the effect of the insertion of fluorophores, in cDNA, on their non radiative and photobleaching decay rates.

The material and experimental set-up are presented in Section 2; Section 3 introduces the theoretical framework; Section 4 is devoted to the experimental results (continuous photobleaching, FCS and sequential photobleaching) and their discussion; the conclusion is made in Section 5.

2. Experimental section

2.1. Material

The experiments have been carried out on two types of single strand cDNA of different size:

i) The so-called long cDNA was prepared from RNA obtained by *in vitro* transcription of the *EcoRI* linearised plasmid pTZ19/PLS-Ta using T7 RNA polymerase (MaxiScript SP6/T7 *in vitro* transcription kit, Ambion).[18] The produced RNA was reverse transcribed in the presence of 500 µM dATP, dGTP and dTTP. For dCTP three different solutions were used. The final concentration of dCTP was kept at 30 µM but contained either 67% (20 µM) or 100% Alexa Fluor 647 labelled dCTP (Invitrogen).

ii) To produce the short cDNA, part of the pTZ19/PLS-Ta sequence was cloned after PCR amplification (primers 5'-CATTAAGGCCTAATTTATGTCG-3' and 5'-AAGCCCGCACTGTCAGG-3') into the vector pCR2.1 (Invitrogen). The resulting plasmid was transcribed with T7 RNA polymerase after linearization with *HindIII* and the obtained RNA was reverse transcribed under the same conditions as for the long cDNA (see above). The sequence of the primer used for cDNA synthesis was 5'-CGTGCCTGTTCTTCGCGTCC-3' for the long and the short cDNA.

Unincorporated nucleotides were separated from the cDNAs by migration in a 1.5 % agarose gel followed by elution of the cDNAs from the agarose gel (High Pure PCR Product Purification Kit, Roche). The long cDNA consists of 306 bases including 74 dCTP deoxynucleotides in its sequence and the short cDNA consists of 147 bases including 45 dCTPs in its sequence.

As shown in Figure 1a the presence of Alexa Fluor 647 labelled deoxynucleotides in the reverse transcription reaction diminishes the quantity of the produced cDNA considerably. No cDNA is detectable if only Alexa Fluor 647 labelled dCTP is present in the reaction. In the following, all cDNA synthesis was performed in the presence of 67% Alexa-Fluor-647-dCTP.

A specific problem that arose with the short cDNA concerned dimerisation. As we see in Figure 1a, the short cDNA migrates at a size that is larger than 200 nucleotides. This band corresponds to cDNA dimers. The two strands can be separated by denaturation of the sample at 95°C for 5 mn before loading (Figure 1b). However, in the following, we will only consider the case of short hybridized (*i.e.* native) cDNA strands, unless specified.

Free Alexa Fluor 647 (Alexa) and Cy5 dyes were purchased from Invitrogen Molecular Probes and stock solution were prepared without further purification.

2.2 Photobleaching set-up

In order to achieve photobleaching of the fluorophores in solution in a reasonable amount of time (a few minutes), with a reasonable laser power (a few mW), the solutions were confined in small volumes. Sample solutions were thus confined into poly(dimethylsiloxane) (PDMS) micro-wells. These micropatterned surfaces were prepared in two steps. First the negative pattern was obtained by standard UV lithography on Si substrate. A regular array containing spots of 60 µm in diameter and 50µm in depth with a pitch of 350 µm was fabricated on the Si wafer. Then, a 10:1 mixture of PDMS base with curing agent (Sylgard 184, Dow Corning) was poured directly on the wafer and baked at 65°C for 2 hours. After peeling off the mold, the 2-mm thick PDMS micropatterned slabs were treated with atmospheric RF plasma to make the surface hydrophilic and facilitate PDMS bonding with glass slides. After this treatment, a drop of the solution of interest (Alexa or cDNA in water)

was spread on the surface. Excess solution was removed and a glass coverslip was placed on top of the PDMS surface to cover the micro-wells. We were control that the wells were filled with solution using a optical microscope in transmission. The experiments were performed within *c.a.* 3 hours following the sample preparation, in order to avoid evaporation.

The experimental setup consisted of a helium-neon laser for excitation (HRP170, Thorlabs) emitting 17 mW maximum power at 633 nm wavelength. This laser beam was sent through a beam expanding telescope and a focusing lens, then into an inverted microscope. In the object plane of the microscope objective (Olympus UMPlanFl, 10×), the laser spot had a Gaussian shape with a 230-μm waist (at $1/e^2$ radius) as measured using a uniformly fluorescent polymer film. This beam was large compared to the micro-wells (diameter 60 μm), so that the excitation intensity was supposed to be constant over a micro-well placed at the center of the laser spot. The excitation intensity was varied from 5 W/cm^2 to 18 W/cm^2. The fluorescence emission from the solution trapped in the micro-well was collected in an epifluorescence configuration by the microscope objective and imaged on a cooled CCD camera (Princeton Instruments). The fluorescence emission signal was obtained by averaging the camera pixels values in a central 26 μm square region. Time traces of the fluorescence signal showed the fluorescence decay due to photobleaching. At each excitation power, the fluorescence decays obtained from three micro-wells were averaged and the photobleaching cross-sections were deduced from a linear fit of the bleaching rate as a function of excitation powers.

2.3 Fluorescence Correlation Spectroscopy set-up

We have used a home made, though rather standard, experimental set-up to record the time trace of the fluorescence fluctuation intensity of the molecules in solution from which we could extract the autocorrelation function as mentioned in the theoretical section. Our set-up is mostly composed of an inverted microscope (Olympus IX70) working in the confocal configuration. the fibered incident laser beam of power 50 to 100 μW (637 nm, LDM635, Thorlabs), was firstly collimated and expanded (×2.4) through a telescope, then sent, *via* a periscope, to a dichroic mirror (z633rdc, Chroma), before being focused within the sample, through the water objective (60×, NA 1.2) of the inverted microscope. The fluorescence light was then spectrally filtered (HQ700/75m, Chroma) and focused on a multimode fiber (core diameter 100 μm), used as point detector. Note that a pair of lenses, located after the tube lens, produce an additional 3× magnification, such that the total magnification was about 180×. Finally, an Avalanche Photodiode, (SPCM-AQR-13, PerkinElmer) was connected to a home made data acquisition system and correlator. Typically a series of 5 acquisitions of 10 to 20 s was performed for each FCS measurement, to properly calculate the averaged autocorrelation curve and standard error of the mean. Data plotting and non linear least square fitting were performed using Origin software (OriginLab Corporation), to extract the number of molecules, N_{eff}, and brightness, ε_{eff}, as explained in the following section.

This set-up has also been used to sequentially photobleach the cDNA solutions, confined into PDMS micro-wells, in combination with FCS measurements. The corresponding protocol consisted in increasing the laser power to several mW, while monitoring the total fluorescence decrease, until it has typically dropped by a factor 1.5. Then the laser power was adjusted back to a power of a few tens of μW, to avoid photobleaching and saturation during the consecutive FCS measurement. When finished, a new photobleaching sequence begins and so on.

3. Theory

3.1 Fluorescence Correlation Spectroscopy

The Fluorescence Correlation Spectroscopy (FCS) is an experimental method that aims at measuring concentration (or number density), diffusion constant and brightness of molecules (or particles) in solution, cellular media, membranes, *etc.* This method exploits the temporal fluctuations of the fluorescence intensity emitted by a small amount of fluorescent molecules diffusing in and out a detection volume, generally under confocal geometry. Using single photon detector like avalanche photodiode, it is possible to detect and measure concentrations from about 0.1 nM to 1 µM. It is thus perfectly adapted to the present situation where we want to analyze cDNA solutions in the nM range, mainly to get information in terms of number concentration and molecular brightness.

For quantitative evaluation, the FCS analysis is performed using the so-called autocorrelation function $G(\tau)$, the amplitude of which is inversely proportional to the effective mean number of molecules in the observation volume, N_{eff} :

$$G(\tau)=\frac{\langle \mathrm{I(t)I(t+\tau)}\rangle}{\langle \mathrm{I(t)}\rangle^2}=1+\frac{\gamma_2}{\mathrm{N_{eff}}}g(\tau) \qquad (1)$$

where γ_2 is a geometric factor that takes into account the exact shape of the observation volume. Using factorial cumulant analysis, we estimated this factor to be 0.23 (see next Section 3.2).

The time dependant part, $g(\tau)$, is characterized by the diffusion time, τ_D, that is dependant upon the diffusion constant, D, and the radius of the confocal observation volume, w_0, according to:

$$\tau_D = w_0^2/4D \qquad (2)$$

Another important output of FCS experiments, that will be used throughout this paper, is the effective photon count rate per molecule, defined as the total photon count rate, CR, divided by the effective number of molecules:

$$CRM_{eff} = CR/N_{eff} \qquad (3)$$

For more details about FCS data analysis see for instance.[16, 19] Throughout this paper we will also refer to the so called effective brightness, ε_{eff}, defined as the number of photon counts per molecule, during a given binning time, δt (that must be much smaller than the diffusion time):

$$\varepsilon_{eff} = CRM_{eff} \times \delta t \qquad (4)$$

The theoretical interest of the brightness will be introduced in the next section. Note that the subscript $_{eff}$ used for the number of molecules, the count rate per molecule or the brightness, comes from the fact that these quantities can be evaluated for a non uniform population of molecules, with a statistical distribution of brightnesses, characterized by its different

moments $\overline{\varepsilon^r}$. In general, ε_{eff} is different from $\bar{\varepsilon}$, so is N_{eff} from the actual number of molecules.

3.2 Factorial cumulants

The factorial cumulants, $\kappa_{[i]}$, are statistical quantities related to the moments of photon counts, k,[20] detected during a given binning time, δt. Up to order 5, the cumulants are defined by:

$$\kappa_{[1]} = \langle k \rangle \tag{5a}$$

$$\kappa_{[2]} = \langle \Delta k^2 \rangle - \langle k \rangle \tag{5b}$$

$$\kappa_{[3]} = \langle \Delta k^3 \rangle - 3\langle \Delta k^2 \rangle + 2\langle k \rangle \tag{5c}$$

$$\kappa_{[4]} = \langle \Delta k^4 \rangle - 6\langle \Delta k^3 \rangle - 3\langle \Delta k^2 \rangle^2 + 11\langle \Delta k^2 \rangle - 6\langle k \rangle \tag{5d}$$

$$\kappa_{[5]} = \langle \Delta k^5 \rangle - 10\langle \Delta k^3 \rangle\langle \Delta k^2 \rangle - 10\langle \Delta k^4 \rangle + 35\langle \Delta k^3 \rangle + 30\langle \Delta k^2 \rangle^2 - 50\langle \Delta k^2 \rangle + 24\langle k \rangle \tag{5e}$$

Following Müller,[17] if the photon counts are those of a population of fluorescent molecules (in our case, cDNA strands) having a distribution of brightnesses characterized by its moments $\overline{\varepsilon^r}$, the factorial cumulants of order r read (for a binning time, δt, much smaller than the diffusion time):

$$\kappa_{[r]} = \gamma_r N_{DNA} \overline{\varepsilon^r} \tag{6}$$

where γ_r is a geometric factor depending upon the shape of the observation volume (γ_2 was introduced in the previous section when defining the autocorrelation function) and N_{DNA} stands for the number of all cDNA molecules in the observation volume (*i.e.* including those having a nil brightness). Note that, correspondingly, the moments $\overline{\varepsilon^r}$ are also defined over the whole population of molecules. In case of one photon excitation, γ_r is well approximated by the formulae $\gamma_r = 1/[(1 + F)r^{3/2}]$, except for γ_1, which is equal to 1.[21] F is a parameter that accounts for the deviation of the observation volume from a perfect Gaussian shape.

If the photon counts proceed from a homogenous solution of N fluorescent molecules in the observation volume, with all the same brightness, ε, the two first cumulants read again:

$$\kappa_{[1]} = N\varepsilon \tag{7a}$$

$$\kappa_{[2]} = \gamma_2 N\varepsilon^2 \tag{7b}$$

Of course, in this case, $\varepsilon = \varepsilon_{eff} = \bar{\varepsilon}$ and $N = N_{eff}$. Consequently, one can calculate the number of molecules, N, with the ratio:

$$N = \gamma_2 \frac{\kappa_{[1]}^2}{\kappa_{[2]}} \tag{8a}$$

and the brightness, ε, from:

$$\varepsilon = \frac{1}{\gamma_2} \frac{\kappa_{[2]}}{\kappa_{[1]}} \tag{8b}$$

Here, we emphasize the fact that calculating N from the two first cumulants is strictly equivalent to calculate it with the autocorrelation function.

The parameter F can be determined from the ratio $\kappa_{[2]}^2/\kappa_{[1]}\kappa_{[3]}$ by using a pure solution of Alexa, because, in this case, this ratio is equal to $\gamma_2^2/\gamma_3 = 3^{3/2}/[8(1 + F)]$. We found $F = 0.56$, that is $\gamma_2 = 0.23$, instead of 0.35 for a perfect gaussian shape.

3.3 Statistical distribution of the brightnesses

If the solution of interest is a mixture of molecules (*i.e*. cDNA strands) having a distribution of brightnesses, characterized by a first and a second moment, $\bar{\varepsilon}$ and $\overline{\varepsilon^2}$, the two first cumulants read:

$$\kappa_{[1]} = N_{DNA}\bar{\varepsilon} \tag{9a}$$

$$\kappa_{[2]} = \gamma_2 N_{DNA}\overline{\varepsilon^2} \tag{9b}$$

One now assumes that the brightness distribution of cDNA molecules is solely due to the distribution of the number, n, of Alexa fluorophores labelling the cDNA strands. Denoting ε_A the individual brightness of the fluorophores, we thus have $\bar{\varepsilon} = \varepsilon_A\bar{n}$ and $\overline{\varepsilon^2} = \varepsilon_A{}^2\overline{n^2}$. For a poissonian distribution of average value $\bar{n}$, the two first cumulants then become:

$$\kappa_{[1]} = N_{DNA}\varepsilon_A\bar{n} \tag{10a}$$

$$\kappa_{[2]} = \gamma_2 N_{DNA}\varepsilon_A{}^2\bar{n}(\bar{n}+1) \tag{10b}$$

This makes it possible to derive the effective number of molecules, N_{eff}, and the effective brightness, ε_{eff}, of the cDNA strands:

$$N_{eff} = \gamma_2 \frac{\kappa_{[1]}^2}{\kappa_{[2]}} = N_{DNA}\frac{\bar{n}}{\bar{n}+1} \tag{11a}$$

$$\varepsilon_{eff} = \frac{1}{\gamma_2}\frac{\kappa_{[2]}}{\kappa_{[1]}} = \varepsilon_A(\bar{n}+1) \tag{11b}$$

Note that, as for a population of uniform brightness, the effective number of molecules, N_{eff} (and their effective brightness, ε_{eff}) can be equally calculated from the autocorrelation function. We point out the fact that N_{eff} is smaller than the total number of molecules, N_{DNA}, because some molecules have a nil brightness. Consistently, ε_{eff} does not tend to zero when $\overline{n}$ does, but to ε_A, because there are always some cDNA strands left that bear, at least, 1 Alexa label.

It is now interesting to use the properties of the fluorophore distribution to relate the effective count rate per molecule and number of molecules, CRM_{eff} and N_{eff}, to the total count rate, CR. With that goal in mind, let us recall that:

$$CR = \varepsilon_{eff} \times N_{eff} / \delta t \tag{12}$$

where δt is the binning time. Using Eq. (11a) and (11b) we thus obtain:

$$CR = \varepsilon_A N_{DNA} \overline{n} / \delta t \tag{13}$$

Consequently, $\overline{n}$ can now be eliminated from Eq. (11) to express the effective number of molecules, N_{eff} and count rate per molecule, CRM_{eff} (equal to $\varepsilon_{eff}/\delta t$), as a function of the total count rate, CR:

$$N_{eff} = N_{DNA} \frac{CR}{N_{DNA} CRM_A + CR} \tag{14a}$$

$$CRM_{eff} = CRM_A \left(1 + \frac{CR}{N_{DNA} CRM_A} \right) \tag{14b}$$

with $CRM_A = \varepsilon_A/\delta t$, the count rate per Alexa fluorophore. Note that the two Eq. (14) are in fact strictly equivalent to each other, since $CRM_{eff} = CR/N_{eff}$. The parameters N_{DNA} and CRM_A can be obtained experimentally, by sequentially photobleaching a cDNA solution: using FCS, the effective photon count rate per molecule and number of molecules are calculated; then one can fit the experimental curves, $CRM_{eff}(CR)$ and $N_{eff}(CR)$, to obtain the values of N_{DNA} and CRM_A. Finally, the average number of labels before photobleaching (*i.e.* when the total count rate is still CR_0) can be calculated with:

$$\overline{n} = \frac{CR_0}{CRM_A N_{DNA}} \tag{15}$$

Let us now consider the consequence of the brightness distribution for the factorial cumulants of higher order (3 to 5). For that purpose we characterize the cumulants through the following ratio:

$$\frac{\kappa[r]^2}{\kappa[r-1]\kappa[r+1]} = \frac{\gamma_r{}^2}{\gamma_{r-1}\gamma_{r+1}} \frac{\left(\overline{\varepsilon^r}\right)^2}{\left(\overline{\varepsilon^{r-1}}\right)\left(\overline{\varepsilon^{r+1}}\right)} \tag{16}$$

The first term on the right hand side of Eq. (16), let us call it $A(r)$, depends upon the observation volume, but is independent of the brightness distribution. The second term on the right hand side is equal to 1 for a solution containing free dyes, the brightness of which is single valued. We can thus measure the first geometrical term $A(r)$ with a solution of Alexa and use it to obtain the normalized ratios of factorial cumulants, $K_{rr,r-1,r+1}$, that depends only of the brightness distribution. Assuming that this latter reflects solely that of the number, n, of Alexa labels in cDNA strands (*i.e.* $\overline{\varepsilon^r} = \varepsilon_A{}^r \overline{n^r}$), we thus obtain:

$$\mathrm{K}_{rr,r-1,r+1} = \frac{1}{A(r)} \frac{\kappa[r]^2}{\kappa[r-1]\kappa[r+1]} = \frac{\left(\overline{n^r}\right)^2}{\left(\overline{n^{r-1}}\right)\left(\overline{n^{r+1}}\right)} \qquad (17)$$

where $\overline{n^r}$ are the different moments of the distribution. If the distribution is poissonian, the moments are all related to the first moment, $\overline{n}$, according to:

$$\overline{n^2} = \overline{n}\left(1+\overline{n}\right) \qquad (18a)$$

$$\overline{n^3} = \overline{n}\left(1+3\overline{n}+\overline{n}^2\right) \qquad (18b)$$

$$\overline{n^4} = \overline{n}\left(1+7\overline{n}+6\overline{n}^2+\overline{n}^3\right) \qquad (18c)$$

$$\overline{n^5} = \overline{n}\left(1+15\overline{n}+25\overline{n}^2+10\overline{n}^3+\overline{n}^4\right) \qquad (18d)$$

In this case, the coefficients $K_{rr,r-1,r+1}$ depend only on the average number of fluorophores per cDNA strand, $\overline{n}$, as can be seen in Figure 2.

Note that the upper limit of $K_{rr,r-1,r+1}$ is 1, that corresponds to a single valued distribution. It is asymptotically attained when $\overline{n}$ tends to infinity, because in this case the relative root mean square of the poissonian distribution tends to zero. It is also attained when $\overline{n}$ tends to zero, because then, the only molecules that are bright are those having 1 Alexa label and no more.

3.4 Photophysical properties of the molecules

Throughout this paper we consider that the laser intensity is sufficiently low so that the molecular transitions are not saturated and that the triplet state population is negligible. In addition, we will assume that the absorption cross-section and the radiative decay rate, k_r^0, of the excited state are the same for free Alexa dyes and individual Alexa labels incorporated in cDNA. In these conditions, the photobleaching cross-section is proportional to the photobleaching quantum yield, *i.e.*, for free Alexa dyes:

$$\sigma_B^0 \propto \eta_B^0 = \frac{k_B^0}{k_r^0 + k_{nr}^0 + k_B^0} \qquad (19a)$$

where k_{nr}^0 is the non radiative decay rate and k_B^0 is the photobleaching decay rate of the excited state of free Alexa. In practice, the photobleaching decay rate, k_B^0, is so small compared to $k_{nr}^0 + k_r^0$, that it does not influence the total decay rate appearing at the denominator (the photobleaching quantum yield for free Alexa, η_B^0, is found to be 2.4×10^{-6}). Consequently, Eq. (19a) can read again:

$$\sigma_B^0 \propto \eta_B^0 = \frac{k_B^0}{k_r^0 + k_{nr}^0} \tag{19b}$$

For the same reason, all the following equations have been simplified by neglecting the contribution of the photobleaching decay rate to the total decay rate.

For Alexa incorporated in cDNA strands, the photobleaching cross-section, σ_B, is:

$$\sigma_B \propto \eta_B = \frac{k_B}{k_r^0 + k_{nr}} \tag{19c}$$

where k_{nr} is the non radiative decay rate, k_B is the photobleaching decay rate and η_B is the phototobleaching quantum yield. It is worthwhile to introduce a normalized photobleaching cross-section, α_B, defined as the ratio of the photobleaching cross-section of the studied species to that of free Alexa:

$$\alpha_B = \frac{\sigma_B}{\sigma_B^0} = \frac{k_r^0 + k_{nr}^0}{k_r^0 + k_{nr}} \times \frac{k_B}{k_B^0} \tag{20}$$

Note that k_B must not be confused with the decay rate of the fluorescence signal recorded by continuously photobleaching a small sample volume:

$$\frac{1}{\tau_B} = \sigma_B I_{ph} = \sigma_B \frac{2P}{\pi w^2} \tag{21}$$

where τ_B is the time constant, I_{ph} is the incident photon intensity (in photon/s/m^2), supposed uniform over the sample, w is the waist of the excitation laser in the sample plane and P is the total power of the excitation laser (in photon/s).

The brightness of the molecules being also proportional to the fluorescence quantum yield, we thus have, for free Alexa dyes in solution:

$$\varepsilon_A^0 \propto \eta_F^0 = \frac{k_r^0}{k_r^0 + k_{nr}^0} \tag{22a}$$

and for individual Alexa labels incorporated in cDNA strands:

$$\varepsilon_A \propto \eta_F = \frac{k_r^0}{k_r^0 + k_{nr}} \tag{22b}$$

For cDNA strands, a first simple reasoning would state that the effective brightness of strands having *n* independent Alexa labels is given by:

$$\varepsilon_{eff} = n\varepsilon_A \propto n\frac{k_r^0}{k_r^0 + k_{nr}} \tag{22c}$$

Consequently, the normalized brightness of the strands, α_ε, is given by:

$$\alpha_\varepsilon = \frac{\varepsilon_{eff}}{\varepsilon_A^0} = n\frac{k_r^0 + k_{nr}^0}{k_r^0 + k_{nr}} \tag{23}$$

It is now worthwhile to consider the ratio $\alpha_\varepsilon/\alpha_B$, because it only depends upon the number of Alexa labels on cDNA strands and on the photobleaching decay rates:

$$\frac{\alpha_\varepsilon}{\alpha_B} = n \times \frac{k_B^0}{k_B} \tag{24}$$

The physical interpretation of this equation is the following: the relative variation of the photobleaching cross-section and of the brightness is only dependent on the number of labels and on the photobleaching decay rate, k_B, but not on k_{nr}. The reason is that when k_{nr} changes, the lifetime of the excited state changes (and hence the fractional time the molecules spend in the excited state), thus multiplying by the same factor the photobleaching cross-section and the brightness.

Another worthwhile equation relates the brightness of free Alexa dyes to that of individual Alexa in cDNA. Using Eq. (24) and the fact that $\varepsilon_{eff} = \alpha_\varepsilon \varepsilon_A^0 = n\varepsilon_A$, we obtain :

$$\varepsilon_A = \varepsilon_A^0 \times \alpha_B \frac{k_B^0}{k_B} \tag{25}$$

4. Results and Discussions

4.1 Continuous photobleaching

Continuous photobleaching experiments were performed with aqueous solutions of Alexa, Alexa-dCTP, and Alexa labelled cDNA strands of both lengths. We checked that the so called photobleaching process is really irreversible, on the time scale of our experiment. First, we want to point out that the fluorescence signal follows a perfectly mono-exponential decay, for all the samples and excitation powers studied. As shown in Figure 3, the time derivative of the fluorescence decay on a semi-log plot is a straight line over more than two orders of magnitude.

For the Alexa dye solution, such an exponential decay was expected. It simply confirms that our measurement is not distorted by other effects such as diffusion inside the sample or adsorption on the walls of the wells.

For the cDNA strands, this mono-exponential behaviour is more surprising, since we expected the cDNA strands to be labelled with numerous Alexa dyes incorporated at various positions along the strand (the long strands possess 74 C bases which may be attached to a dye), that might undergo some self-quenching [22, 23], due to Förster-type energy homo-transfer. If self-quenching occurred, the excited state lifetime of individual fluorescent labels (hence their brightness and the photobleaching decay rate) is expected to vary during photobleaching, since the labelling density varies. The mono-exponential decay observed for both cDNA samples suggests that the Alexa labels of the same strand behave like independent fluorophores with negligible interactions between each other, though they may interact with the cDNA backbone, as we shall see.

From data shown in Figure 3, the photobleaching rate could be obtained for Alexa dye, Alexa-dCTP and cDNA strands. We see that the addition of the dCTP side group to the Alexa dye only slightly affects its photobleaching rate, whereas the cDNA-substituted Alexa dye exhibits a noticeably slower photobleaching rate. The fluorescence decay time constant, τ_B, of Alexa linked to cDNA is longer by approximately a factor of two compared to that of free Alexa, at the same excitation intensity.

We performed fluorescence decay measurements at several incident laser powers in the 4 mW to 15 mW range. The fluorescence decay rate, given by $1/\tau_B$, is measured on the mono-exponential decay curves, the laser intensity being supposed uniform over the surface of the micro-well containing the sample solution. As expected, it depends linearly on the incident power for all samples and is related to the photobleaching cross-section, σ_B, through Eq. (21).

Table 5 summarizes the photobleaching cross-section values for the four samples. In addition, in the case of Alexa dye, we could determine the photobleaching quantum yield which is the ratio between the photobleaching and absorption cross-sections. It represents the probability of the dye to bleach for one adsorption/relaxation cycle. The absorption cross-section was obtained independently from transmission measurement on concentrated solutions. We found the photobleaching quantum yield for Alexa to be 2.4×10^{-6}. As a comparison, the same measurements were performed with Cy5 dye and yielded a photobleaching quantum yield of 4.7×10^{-6} for Cy5, is in good agreement with the value of 5×10^{-6} determined elsewhere[24]. These results suggest that Alexa is approximately twice more photostable than Cy5, which is also consistent with the literature. [25]

Although substitution with the dCTP side-group has little effect on the bleaching rate of Alexa, the Alexa dye when linked to cDNA is more photostable by a factor of two compared to free Alexa. Assuming that the absorption cross-section is independent of the

environment of the dye, there are two different interpretations for this reduced photobleaching cross-section. One possible reason is an increase of the non-radiative relaxation rate (k_{nr}) due to quenching of the Alexa dye by the proximity of the cDNA backbone (the self quenching could have the same effect [1], but this phenomenon is ruled out in our case from the observed monoexponential decay discussed above). In this case, the excited state lifetime would decrease, which would lead to a reduced probability of photobleaching per unit of time, since this occurs when the dyes are in the excited state. A second possible explanation is simply a reduced photobleaching rate, k_B, for excited Alexa when linked to cDNA strands, either because it maintains a more rigid conformation or because it is shielded from reactive molecules by the cDNA.

4.2 Fluorescence Correlation Spectroscopy

4.2.1 Diffusion time, concentration and brightness measurements

By fitting the autocorrelation function to the appropriate model for free diffusion the effective numbers of molecules in the observation volume N_{eff} and the diffusion time τ_D of the different species (Alexa, Alexa-dCTP, short and long cDNA) can be extracted (Figure 4).

Figure 4a shows that for a given sample divided in two sub-samples, the number of short cDNA strands from the denatured sub-sample is about twice that of the short cDNA strands from the native sub-sample. This observation is consistent with the agarose gel analyze (see Section 2.1), that showed that short cDNA strands hybridize and form dimers.

Figure 4b (and results displayed in Table 5) shows that cDNA strands diffuse about ten times slower than Alexa fluorophores. Of course, because of their much larger molecular mass cDNA molecules have an hydrodynamic radius much larger than Alexa fluorophores. However a precise theoretical determination of their hydrodynamic radius depends upon their sequence and the pH, salinity of the solution, *etc*.[26]

Actually, in the scope of the present paper, the most important data that we could extract from the FCS measurement is the effective photon count rate per molecule, CRM_{eff} (Table 5), or, equivalently, the effective brightness of the cDNA strands, ε_{eff}. We thus found that the effective brightness of the long strands and of the short dimer strands is about twice that of Alexa, while that of the denaturated short cDNA strands is very close to Alexa.

So far, we cannot unravel the contributions, for the cDNA brightness, of the number of Alexa fluorophores per cDNA strand and of their individual brightness: there can be, either a large number of fluorophores per strand with a low brightness, or a few fluorophores with a brightness very similar to that of free Alexa. In the next section, we will see that one can make profit of the photobleaching and FCS experiments to solve this indistinctness.

4.2.2 cDNA labelling – Model 0

We now want to make a crude estimation of the average number of Alexa fluorophores on cDNA strands from the continuous photobleaching and FCS experiments. For that purpose, we consider the values of the normalized photobleaching cross-sections α_B and of the normalized brightness α_ε (see Table 5).

Let us assume that, in cDNA, their values are solely due to the variation of the radiationless decay rate (k_{nr}) and not to the change of the photobleaching decay rate, k_B.

Therefore, according to Eq. (24), the ratio $\alpha_\varepsilon / \alpha_B$ gives the number, n, of Alexa labels on the strands. We see, in Table 5, that α_B is 0.49 and 0.53 for short and long cDNA strands and that α_ε = 1.89 and 2.05 for short and long cDNA strands. This immediately provides a first crude estimation of the number of Alexa fluophores per cDNA strands: n = 3.86 for short cDNA strands and n = 3.87 for long cDNA strands.

In addition, it is interesting to observe that Alexa-dCTP photobleaches slightly faster than Alexa (α_B = 1.07), in agreement with its slightly higher brightness (α_ε = 1.07): this may reflect the slightly reduced non radiative decay rate, k_{nr}, for Alexa labels incorporated in cDNA.

We could have also made a second crude estimation of the average number of Alexa fluorophores on cDNA strands, by considering that only the photobleaching decay rate, k_B, of Alexa fluorophores changes in presence of the cDNA backbone, while the non-radiative decay rate, k_{nr}, is not affected. Thus, according to Eq. (23), the normalized brightness, α_ε, directly gives the number of Alexa labels: n = 1.89 and 2.05 for short and long cDNA strands.

4.3 Statistical distribution of the number of fluorescent labels

4.3.1 Statistical distribution of the brightnesses

The above estimated numbers of labeled C-bases being much smaller than the number of C bases of the considered sequences of cDNA (74 C bases on long cDNA strands and 2×45 on hybridized short cDNA strands), it is reasonable to assume that the number, n, of labeled C-bases must be distributed and that the molecular brightness has statistical properties that reflects this distribution. In addition, we consider that this distribution is the only phenomenon at the origin of the brightness fluctuations and that it follows a Poisson law. Since the next sections of the present paper are based on this assumption, we want to provide a direct, although non very accurate, proof of this idea. For that purpose we used the normalized ratios of factorial cumulants, $K_{rr,r-1,r+1}$, introduced in the theoretical Section 3.3 [see Eq. (17)]. One can compare the experimental values of the ratios $K_{rr,r-1,r+1}$ to the theoretical ones, shown in Figure 2.

Due to the uncertainty of the experimental results, that increases with r, the useful ratios $K_{rr,r-1,r+1}$ are limited to 3 and 4. We exclude r = 2 from the discussion, because we observed that the ratio $K_{22,1,3}$ is always below the values expected for a poissonian distribution. We attribute this finding to the fact that $\kappa_{[1]}$ can be increased by scattered light and other parasitical photons, while higher order cumulants cannot (because the corresponding photons have a poissonian distribution). $K_{22,1,3}$ being inversely proportional to the first order factorial cumulant, the parasitical photons decrease the value of the ratio $K_{22,1,3}$.

As an experimental result, we systematically observed that the values $K_{rr,r-1,r+1}$ were significantly smaller than 1, the value corresponding to, either a very small number of labels or a very large one, or, more generally, to a single valued distribution. We found $K_{33,2,4}$ = 0.88 ± 0.08 and $K_{44,3,5}$ = 0.9 ± 0.1. Despite a significant uncertainty and variability, this definitively proves that our cDNA strands are not uniformly labeled and that $\overline{n}$ is very probably limited to a small value.

Part of the experimental variability comes from the fact that the long cDNA strands tend to exhibit a lower value of the ratio $K_{rr,r-1,r+1}$ (compared to the short cDNAs). This may be attributed to the contribution of the brightness distribution of the Alexa fluorophores when incorporated in cDNA, a phenomenon that has been neglected in the theoretical calculation

and that can be different for long and short cDNAs. For instance, we checked that a gaussian distribution of the individual fluorophore brightness, ε_A, with a relative root mean square of 0.5, could explain the experimental variability between the long and short cDNA strands. However we will consider this phenomenon as a second order one, compared to the role played by the distribution of the number, *n*, of fluorophores per cDNA strand.

4.3.2 cDNA labelling – Model 1

In the Section 4.2.2, we proposed a zero order model for the cDNA labelling (*Model 0*). Implicitly, we made two assumptions about the brightness, ε_A, of the individual Alexa fluorophores labelling the cDNA. The first assumption followed from the hypothesis of a photobleaching decay rate identical for free Alexa dyes and Alexa labels in cDNA (while the non radiative decay rate, k_{nr}, is modified). Rewriting the Eq. (25) with $k_B = k_B^0$, this leads to:

$$\varepsilon_A = \alpha_B \varepsilon_A^0 \tag{26a}$$

By combining this equation with Eq. (11) [$\varepsilon_{eff} = \varepsilon_A(\overline{n}+1)$], we can thus relate the effective brightness of cDNA strands to that of free Alexa:

$$\varepsilon_{eff} = \alpha_B \varepsilon_A^0 (\overline{n}+1) \tag{26b}$$

In addition, knowing that α_ε is the proportionality factor between the effective brightness of cDNA strands, ε_{eff}, and that of free Alexa, ε_A^0, we thus obtain :

$$\overline{n}+1 = \alpha_\varepsilon / \alpha_B \tag{26c}$$

By comparing this equation to Eq. (24) with $k_B = k_B^0$, we see that the difference is due to the statistical distribution of the number of labels, which is so that *n* is replaced by $\overline{n}+1$. Since $\alpha_\varepsilon/\alpha_B$ is small, this makes a significant difference and the average number of Alexa fluorophores per short and long cDNA strand, $\overline{n}$, is now 2.86 and 2.87, instead of 3.86 and 3.87, as obtained with our first crude estimation (*Model-0*).

We could have also used the second implicit assumption about the brightness of the individual Alexa fluorophores, that followed from the hypothesis that only the photobleaching decay rate, k_B, of Alexa fluorophores changes in presence of the cDNA backbone (while k_{nr} is unaffected by cDNA). By comparing Eq. (22a) and Eq. (22b), we immediately see that this corresponds to:

$$\varepsilon_A = \varepsilon_A^0 \tag{27a}$$

In this case, still using Eq. (11) [$\varepsilon_{eff} = \varepsilon_A(\overline{n}+1)$], we would have:

$$\overline{n}+1 = \alpha_\varepsilon \tag{27b}$$

and thus an average value of the number of Alexa fluorophores per short and long cDNA strand which is $\overline{n}$ = 0.89 and 1.05, instead of 1.89 and 2.05, as obtained with our second crude estimation (*Model-0*).

4.4 Sequential photobleaching experiments

4.4.1 cDNA labelling – Model 2

In the previous Section 4.3.2, we were left with an ambiguity concerning the estimation of the average number of Alexa fluorophores per cDNA strand: about 3 or 1, depending of the hypothesis about the radiationless decay rate (k_{nr}) and photobleaching rate (k_B) of Alexa fluorophores incorporated in cDNA. Hereafter, we provide an estimation that does not require additional hypothesis, other than the assumption that the statistical distribution of fluorophores follows a Poisson law. For this, we will experimentally apply the considerations developed in Section 3.3, where we have shown that the number of fluorescent labels can be deduced from the variations of the effective number of molecules, N_{eff} and of the effective count rate per molecule, CRM_{eff} (or equivalently the brightness, ε_{eff}), during sequential photobleaching of cDNA solutions (see Eqs. 14a and 14b).

This is exemplified in Figs. 5, where we can see, *versus* the total count rate, CR: i) the linear decrease of the effective count rate per molecule, CRM_{eff}, the slope of which is equal to $1/N_{DNA}$ where N_{DNA} denotes the effective number of DNA molecules (fluorescent or not) inside the confocal volume and the intercept of which for $CR = 0$ is equal to CRM_A, the count rate per fluorophore [Eq. (14b)]; ii) the effective number of molecules, N_{eff}, the behavior of which is perfectly predicted by Eq. (14a), using the parameters extracted from the curve $CRM_{eff}(CR)$. Eventually, from these parameters, Eq. (15) gives the mean value of the Poisson law describing the distribution of number of fluorophores: $\overline{n}$ = 1.80 ± 0.22 for the long cDNA strand (Figure 6a is a typical example) and $\overline{n}$ = 1.49 ± 0.46 for the short cDNAs. We also find $\overline{n}$ = 0.49 for the denaturated short cDNAs (see an example in Figure 6b). It is worthwhile to note that $\overline{n}$ is a mean value calculated over the population of all the cDNA strands, *i.e.* including the dark cDNAs, having no Alexa fluorophores. Therefore, the visible cDNAs have a mean number of Alexa labels slightly larger, since it is given by $\overline{\overline{n}} = \overline{n}/(1 - e^{-\overline{n}})$.

Here follows some comments concerning the shape of the experimental curves shown in Figure 6. Let us first consider the limiting case of a large number of fluorophores per cDNA strand: as photobleaching begins to occur, the effective count rate per molecule decreases because cDNA strands are loosing fluorophores, but since no cDNA strand (initially bright) totally photobleaches (because the strands are still carrying numerous fluorophores), the effective number of molecules remains almost constant (Eq. (11) with $\overline{n} >> 1$). Conversely, another limiting case occurs for a population of cDNA strands labeled with such a small number of fluorophores, that the probability to find a strand with two labels is negligible (compared to the probability to find a strand with one label). In that case, the effective count rate per molecule is constant (it is that of cDNA strands with one label), while the effective number of (visible) molecules decreases (Eq. (11) with $\overline{n} << 1$). Eq. (14) exemplify these two limiting cases of the dependency of CRM_{eff} and N_{eff} *versus* CR. The key parameter, that drives the transition between the two limits, is the number of fluorophores per cDNA strand, *i.e.* the ratio $CR/N_{DNA}CRM_A$: when it is large, N_{eff} is almost constant (and equal to N_{DNA}) while CRM_{eff} becomes almost proportional to CR; conversely, when the number of fluorophores per cDNA strand is small, CRM_{eff} is almost constant (and equal to CRM_A), while N_{eff} becomes proportional to CR. These behaviors are put into evidence in Figure 6: i) the

linear decrease, with *CR*, of the effective count rate per molecule, CRM_{eff}, is either significant for the long strands (Figure 6a), or much more limited for the denaturated short strands labeled with a smaller number of fluorophores (Figure 6b); ii) the effective number of molecules, N_{eff}, which either tends to be constant when the long strands are labeled with a large number of fluorophores (Figure 6a at large count rate, *CR*) or becomes proportional to the count rate, *CR*, when there are only a few fluorophores left on the denaturated short strands (Figure 6b).

4.4.2 From cDNA labelling to photophysical properties

To conclude the Results and Discussion section, we want to stress that the average number of fluorophores per short or long cDNA strand deduced from the latter *Model-2* ($\overline{n}$ ~ 1.5 - 1.9), is significantly smaller than the mean value, $\overline{n}$ = 2.86, found by assuming a constant photobleaching rate k_B. Conversely, it is larger than the estimation, $\overline{n}$ = 0.97, made by assuming a constant non radiative decay rate k_{nr}. It is thus reasonable to suggest that the truth is in between these two limiting cases! Indeed we now show that it is possible to compare the results of the sequential photobleaching experiment with those of the continuous photobleaching and FCS experiments (see Sections 4.1 and 4.2.1) to get information about the dependency of the bleaching rate k_B of the excited state upon the environment. As a matter of fact, by combining Eq. (11b) and (25) and remembering that $\varepsilon_{eff} = \alpha_\varepsilon \varepsilon_A^0$, we obtain:

$$\frac{k_B}{k_B^0} = \frac{\alpha_B(\overline{n}+1)}{\alpha_\varepsilon} \qquad (28)$$

Using the values displayed in Table 5 for α_B and α_ε and the values estimated in the Section 4.4.1 for $\overline{n}$, we find k_B / k_B^0 = 0.72 for long cDNA and 0.65 for short non-denatured cDNA. This result suggests that the photobleaching decay rate, k_B, is lower for fluorophores in cDNAs. This is not inconsistent with observations made by Widengren and Schwille. [27] According to these authors who characterized, by FCS, the Cyanine dye Cy5 (from which Alexa is derived), the rates of intersystem crossing and triplet state decays were noticeably reduced for the DNA substituted dye. This fact may reflect that the dye is shielded by the cDNA strand from the oxygen molecules of the aqueous solution.

On the other hand, inserting the values obtained for k_B / k_B^0 in Eq. (20), we can find that the ratio of the decay rates of the excited state, $(k_r^0 + k_{nr}) / (k_r^0 + k_{nr}^0)$, is 1.36 for long cDNA and 1.33 for short non-denatured cDNA. This shows that the non radiative decay rate of Alexa is increased when incorporated in cDNA. Then, according to Eq. (22), the individual brightness of the fluorophores is reduced by the same factor, that is: $\varepsilon_A = 0.74 \varepsilon_A^0$ for long cDNA and $\varepsilon_A = 0.75 \varepsilon_A^0$ for short non-denatured cDNA.

5. Conclusion

We have introduced a theoretical framework to analyse cDNA solutions in terms of multiple labelling. The last approach based on the variation of the number of molecules and brightness *versus* the total count rate during sequential photobleaching (*Model 2*), is independent of any assumption about the photophysical properties of the excited state of the fluorophores (apart from the fact that we have neglected non linear processes and singlet-triplet transitions, because of the moderate power of our laser beam). This model just relies on the hypothesis that the number of labels follows a poissonian distribution, which is justified by the much greater number of available label sites than the degree of labelling.[13] Such a distribution has the important property to be conserved during photobleaching.

When combining our approach with independent and careful photobleaching cross-section and FCS measurements, one gets additional information about the interplay between the radiative, non radiative and photobleaching decay rates in the electronic excited state of the labelling dye. These photophysical outputs are usually provided by dedicated single molecule experiments,[28, 29, 30] or time resolved ones.[31, 32, 33]

Concerning the central question addressed in this paper, the determination of the number of Alexa fluorophores labelling cDNA strands, it may be, at a first glance, surprising to discover that the incorporation during reverse transcription is so inefficient: roughly, only 2.5% of Alexa labelled dCTP were incorporated in the short or long cDNA strands! In fact, it is known that base linked fluorescent nucleotides are poorly incorporated in consecutive position by DNA polymerase.[5, 34, 35] The consequence of this, beside the difficulty to detect weakly labelled strands (when working in the single molecule regime), is that their brightness necessarily fluctuates quite a lot, together with the fact that a significant amount of strands are not labelled at all! More precisely, a typical mean value of 2 labels per cDNA strand corresponds to 13.5% strands without any label; 27% strands labelled with 1 dye, 27% with 2 dyes; 18% with 3 dyes; 9% with 4 dyes; 3.6% with 5 dyes (and 2% of strands labelled with more dyes). Since the brightness of a strand is proportional to the number of dyes it bears (at least at low labelling degrees), all the species with 1 to 5 dyes contribute in a comparable way to the total fluorescence signal.

We stress the fact that the method presented in this paper can be applied to DNA samples of low concentration, while standard measurements of labelling efficiency are based on absorption spectroscopy and so, do require amplification of the genetic material.[35]

The quantification of multiple labelling is a very general problem for protein or DNA analysis involved in biochemistry, pharmacology, medicine, *etc.* The question is especially crucial with the development of ultra sensitive techniques to detect biomolecules, among which, single molecule measurements play a central role but are also, unfortunately, quite sensitive to multiple labelling. We believe that our approach is very general and relatively easy to implement, since it implies bulk measurements with concentrations in the nM - µM range and no single molecule immobilisation and imaging.

6. Acknowledgements

This project was funded by the French Agence Nationale de la Recherche under contract ANR-07-GPLA-013-001.

The authors also gratefully thank Jie Gao for helpful and fine technical assistance on the FCS experiment.

References

(1) R. Luchowski, E. G. Matveeva, I. Gryczynski, E. A. Terpetschnig, L. Patsenker, G. Laczko, J. Borejdo, Z. Gryczynski, *Current Pharmaceutical Biotechnology* **2008**, *9*, 411-420

(2) N.C.Shaner, P.A.Steinbach, R.Y.Tsien, *Nat. Meth.* **2005**, *2*, 905-909

(3) D. Singh, V. Kumer, K.N. Ganesh, *Nucleic Acids Res.* **1990**, *18*, 3339-3345

(4) A.C. Cuppoletti, Y. Cho, J.-S. Park, C. Strässer, E. T. Kool, *Bioconjugate Chem.* **2005**, *15*, 528-534

(5) J. Hesse, J. Jacak, M. Kasper, G. Regl, T. Eichberger, M. Winklmayr, F. Aberger, M. Sonnleitner, R. Schlapak, S. Howorka, L. Muresan, A.-M. Frischauf, G.J. Schütz, *Genome Res.* **2006**, *16*, 1041-1045

(6) M.L. Metzker, *Genome Res.* **2005**, *15*, 1767-1776

(7) H.-A. Wagenknecht, *Ann. N.Y. Acad. Sci.* **2008**, *1130*, 122-130

(8) W. E. Moerner, *J. Phys. Chem. B* **2002**, *106*, 910-927

(9) V.T. Nguyen, Y. Kamio, H. Higuchi, *EMBO J.* **2003**, *22*, 4968-4979

(10) Y.P. Wang, J. Biernat, M. Pickhardt, E.Mandelkow, E.-M. Mandelkow, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 10252-10257

(11) E.H. Koo, P. T. Lansbury Jr., J.W. Kelly, *Proc. Natl. Acad. Sci. USA* **1999**, 96, 9989-9990

(12) S.K. Das, M. Darshi, S. Cheley, M.I. Wallace, H. Bayley, *ChemBioChem*, **2007**, *8*, 994-999

(13) T.C. Messina, H. Kim, J.T. Giurleo, D.S. Talaga, *J. Phys. Chem. B* **2006,** *110,* 16366-16376

(14) S.A. Mutch, B.S. Fujimoto, C.L. Kuyper, J.S. Kuo, S.M. Bajjalieh, D.T. Chiu, *Biophys. J.* **2007**, *92*, 2926-2943

(15) Y. Chen, J. Müller, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 3147-3152, and references therein

(16) E.L. Elson, *J. Biomed. Opt.* **2004**, *9*, 857-864

(17) J. D. Müller, *Biophys. J.* **2004**, *86*, 3981-3992

(18) L.-J. Chen, E. M. Orozco, *Nucleic Acids Res.* **1988**, *16*, 8411-8431

(19) N.L.Thomson *Fluorescence correlation spectroscopy in Topics in Fluorescence Spectroscopy*, (Eds.: J. R. Lakowicz), Plenum Press, New York, **1991**, pp.337–378.

(20) M.G. Kendall, A. Stuart. *Moments and cumulants* in *The advanced theory of statistics*, *Vol 1*, MacMillan Publishing, New York, **1977**, p.57

(21) B. Huang, T. D. Perroud, and R. N. Zare, *ChemPhysChem* **2004**, *5*, 1523-1531

(22) J. E. Berlier, A. Rothe, G. Buller, J. Bradford, D. R. Gray, B. J. Filanoski, W. G. Telford, S. Yue, J. Liu, C.-Y. Cheung, W. Chang, J. D. Hirsch D., J. M. Beechem, R. P. Haugland, R. P. Haugland, *J. Histochem. Cytochem.* **2003**, *51*, 1699-1712

(23) J. R. Lakowicz, J. Malicka, S. D'Auria, I. Gryczynski, *Anal. Biochem.* **2003**, *320*, 13-20

(24) C. Eggeling, J. Widengren, R. Rigler, C. A. M. Seide in *Applied fluorescence in chemistry, biology and medicine*, (Eds.: W. Rettig, B. Strehmel, S. Schrader, H. Seifert), Springer-Verlag, Berlin, **1999**, p. 193

(25) J.E.Berlier, A.Rothe, G.Buller, J.Bradford, D.R.Gray, B.J.Filanoski, W.G. Telford, S.Yue, J.Liu, C.-Y.Cheung, W.Chang, J.D.Hirsch, J.M. Beechem, R.P.Haugland, R.P. Haugland, *J. Histochem. Cytopchem*. **2003**, 51, 1699-1712

(26) B. Tinland, A. Pluen, J. Sturm, G. Weill, *Macromolecules* **1997,** *30*, 5763-5765

(27) J.Widengren, P. Schwille, *J. Phys. Chem. A.* **2000**, *104*, 6416-6428

(28) L.A. Deschenes, D.A. Vanden Bout, *Chem. Phys. Let.* **2002**, *365*, 387-395

(29) E. Füreder-Kitzmüller, J. Hesse, A. Ebner, H.J. Gruber, G.J. Schütz, *Chem. Phys. Let.* **2005**, 404, 13-18
(30) Z. Huang, D. Ji, S. Wang, A. Xia, F. Koberling, M. Patting, R. Erdmann, *J. Phys. Chem. A* **2006,** *110,* 45-50
(31) M.E. Sanborn, B.K. Connolly, K. Gurunathan, M. Levitus, *J. Phys. Chem. B* **2007,** *111,* 11064-11074
(32) P. Sarkar, S.V. Koushik, S.S. Vogel, I. Gryczynski, Z. Gryczynski, *J. Biomed Opt.* **2009**, 14, 340471-340479
(33) I. Gregor, J. Enderlein, *Photochem. Photobiol. Sci.* **2007**, *6*, 13–18
(34) T. Margaritis, P. Lijnzaad1, D. van Leenen, D. Bouwmeester, P. Kemmeren, S.R. van Hooff, F.CP Holstege, *Molecular Systems Biology* **2009,** *5*:266
(35) P.A. C. 't Hoen, F. de Kort, G.J.B. van Ommen, J.T. den Dunnen, *Nucleic Acids Res.* **2003**, *31*, e20

Figure captions and Table

Figure 1. a) Diminution of cDNA production in the presence of Alexa Fluor 647 labelled dCTP (short cDNA strand). The reactants contained no Alexa-Fluor-647-dCTP (lane 2), 20.1 µM Alexa-Fluor-647-dCTP (lane 3) or 30 µM Alexa-Fluor-647-dCTP (lane 4). 1 µl of the 30 µl reverse transcription assays of the short cDNA have been loaded on a 1.5 % agarose gel for analyses. Lane 1 corresponds to molecular size markers. b) Dimerisation of the short cDNA strands. 1µl of the reverse transcription reaction has been analyzed immediately after cDNA synthesis (lane 3) or after denaturation at 95°C for 5 min (lane 1) on a 1.5 % agarose gel. Lane 2 shows molecular size markers.

Figure 2. Normalized ratios of factorial cumulants, $K_{rr,r-1,r+1}$, *versus* $\overline{n}$; the solid, dashed and dotted lines correspond to the theoretical calculations done by assuming a poissonian distribution of the number, n, of Alexa fluorophores per cDNA strand with, respectively, r = 2, 3 and 4.

Figure 3. Time-derivative of the decay of the fluorescence signal, $F(t)$, due to photobleaching, on a semi-log scale measured with 4 mW excitation power. Results are shown for three samples: Alexa (open circles), Alexa-dCTP (open squares) and Alexa labeled long cDNA strands (open triangles). Fluorescence decays obtained from short cDNA strands are similar to that of long strands (data not shown).

Figure 4 Example of experimental autocorrelation functions (symbols), acquired with different species and concentrations, with their fits for free diffusion (solid lines). a) autocorrelation function of short cDNA strands, the amplitude of which is about twice when the strands are hybridized (open circles *versus* open triangles). b) normalized autocorrelation function for Alexa (open triangles) and long cDNA strands (open circles), exemplifying the difference of diffusion times.

Table 5 Mean values of the photobleaching cross-section, σ_B, of the photon count rate per molecule (normalized at 1 µW of excitation laser), of the diffusion time, τ_D and of the normalized photobleaching cross-section and brightness, for Alexa, Alexa-dCTP and cDNA strands.

	Photobleaching cross-section σ_B ($10^{-22}cm^2$)	*Photon count rate per molecule at 1 µW (kHz)*	*Diffusion time* τ_D *(µs)*	$\alpha_B = \sigma_B / \sigma^0_{B,A}$	$\alpha_\varepsilon = \varepsilon_{eff} / \varepsilon^0_A$
Alexa	10.6 ± 0.2	9.5 ± 0.3	59 ± 2	1	1
Alexa-dCTP	11.4 ± 0.2	10.2 ± 0.3	68 ± 3	1.07	1.07
Long cDNA	5.7 ± 0.1	19.4± 2	632 ± 60	0.53	2.05
Short cDNA	5.2 ± 0.1	18.0 ± 2	490 ± 30	0.49	1.89
Denatured short cDNA		11.2	410		1.18

Figure 6 Typical count rate per molecule (CRM_{eff}, solid squares) and effective number of molecules (N_{eff}, open circles) as a function of the total count rate (CR) determined from FCS measurements during photobleaching: a) Long cDNA strands; b) Denatured short cDNA strands. The curves are fit with equations (14) and the initial number of labels per cDNA molecule ($\overline{n}$) is then determined from Eq. (15).

Figure 1

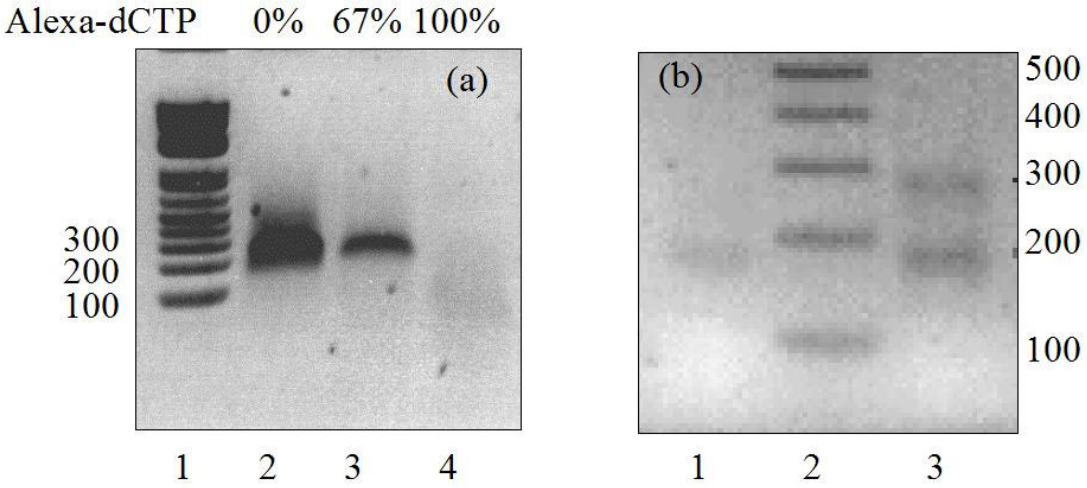


Figure 2

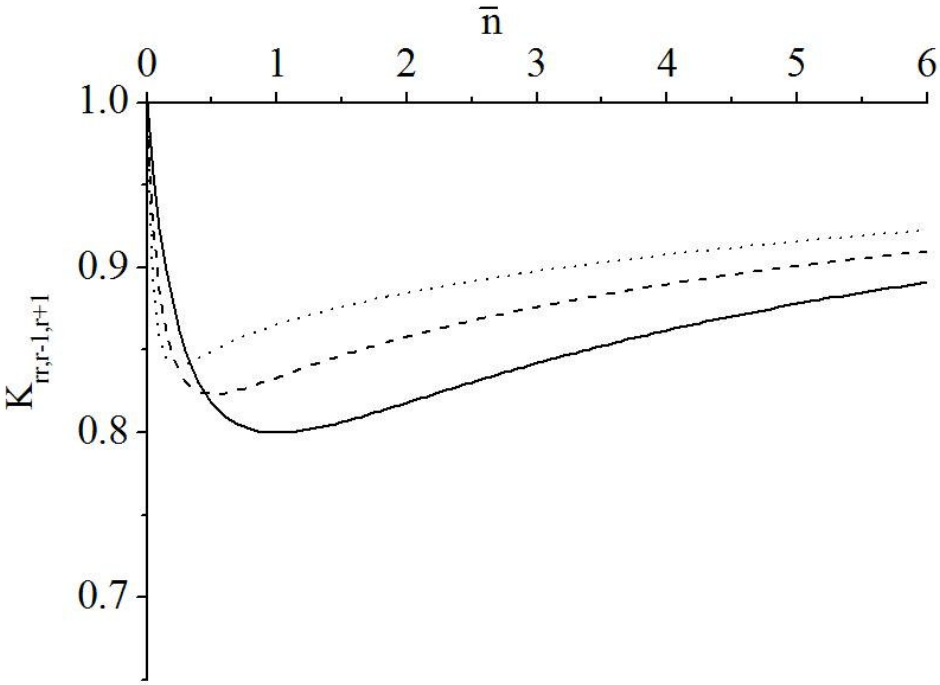


Figure 3

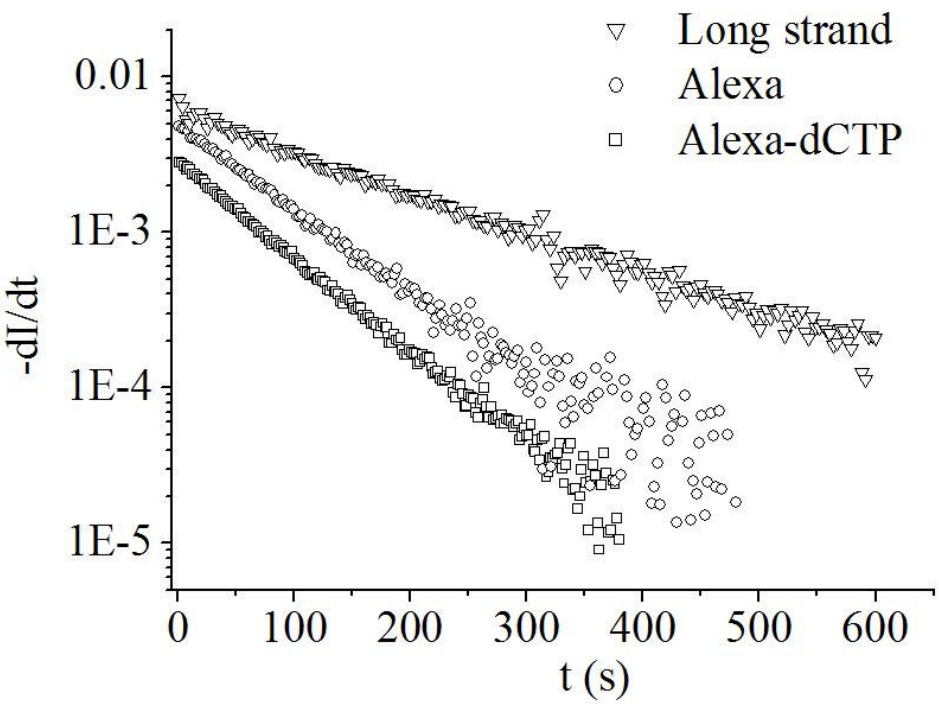


Figure 4

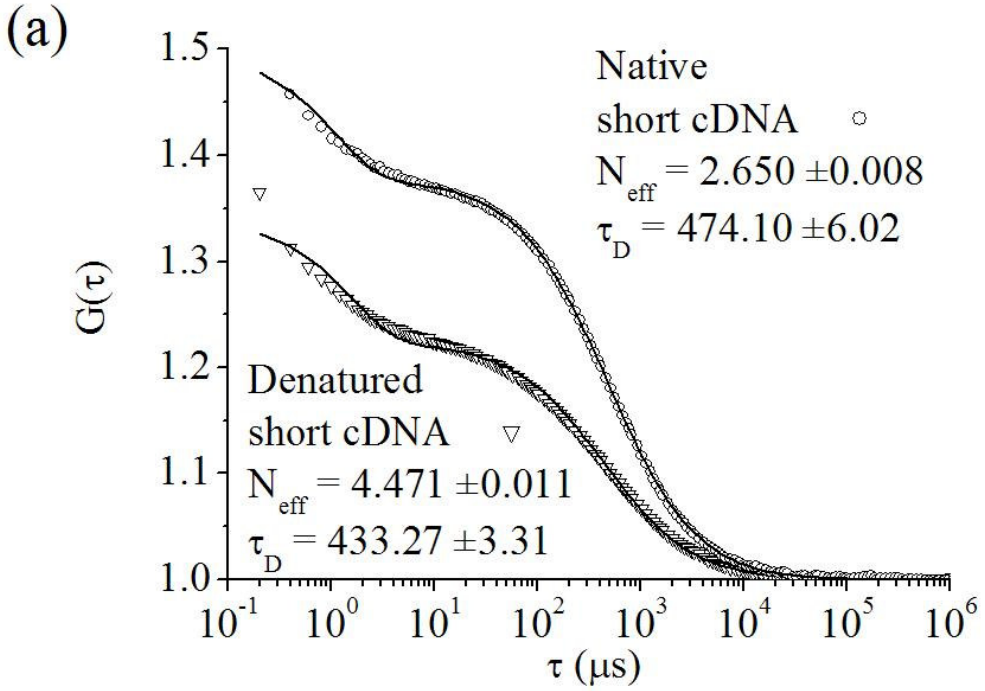


Figure 4

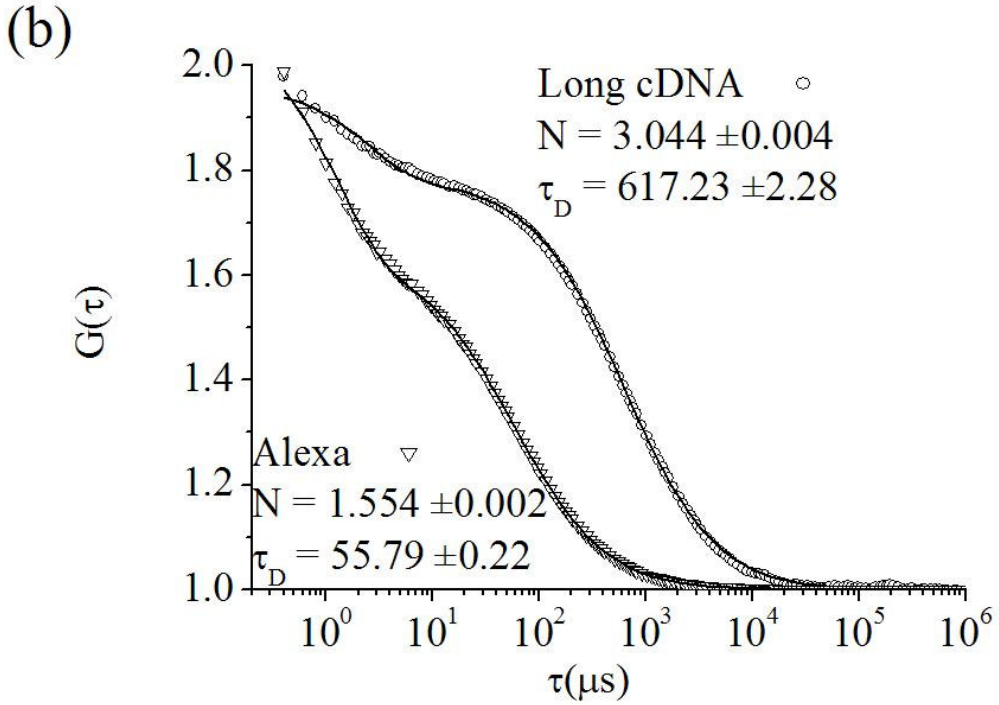


Figure 6

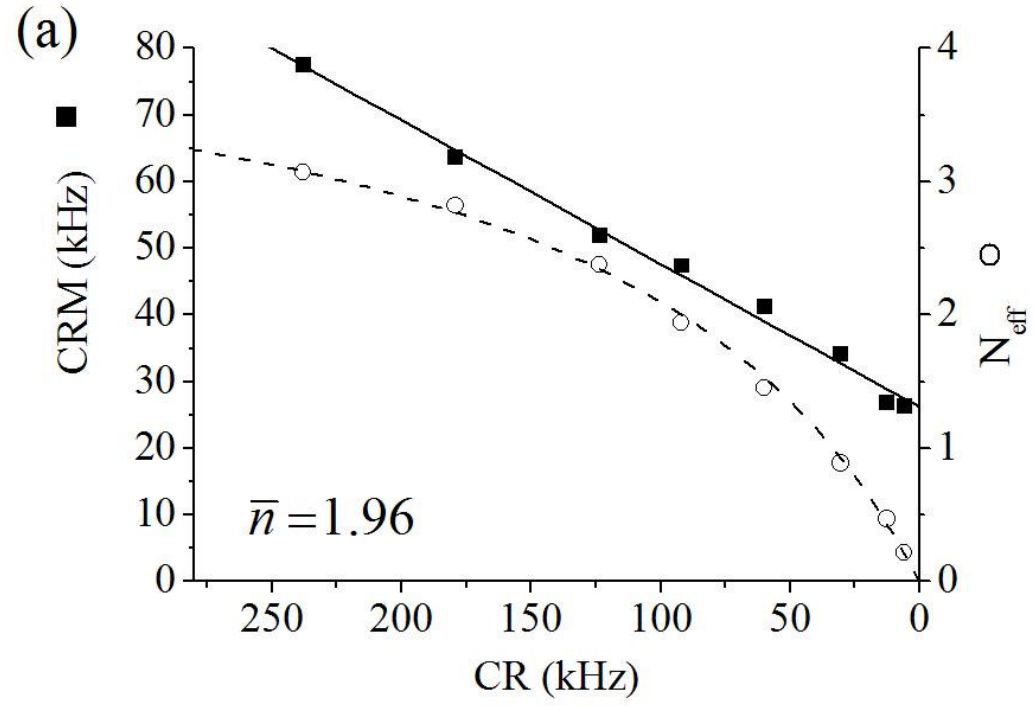
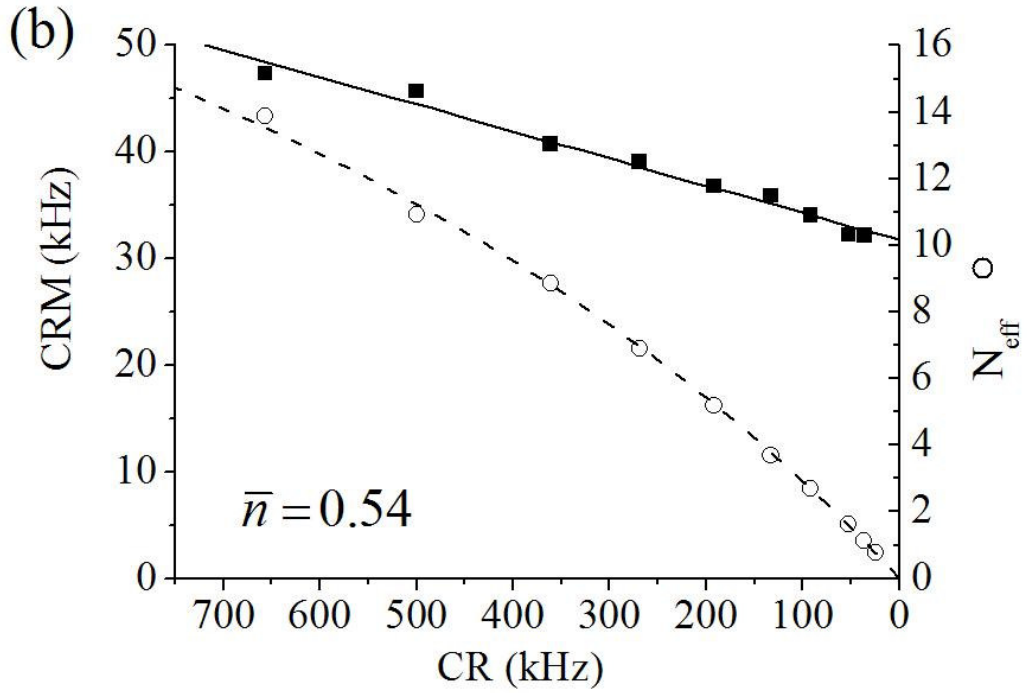


Figure 6


(b)
CRM (kHz)
CR (kHz)
N_{eff}
$\overline{n} = 0.54$
50
40
30
20
10
0
700
600
500
400
300
200
100
0
16
14
12
10
8
6
4
2
0