Fluorescence Correlation Spectroscopy: an Experimentalist's View of the Basics

G. Jung[1]

Biophysical Chemistry, Department 8.1 Chemistry, Saarland University, Im Stadtwald, Building B2.2, 66123 Saarbruecken, Germany

Although Fluorescence Correlation Spectroscopy was developed more than 30 years ago, increasing interest only emerged in the past decade. Especially the implementation in commercial confocal microscopes promoted its spreading in the life sciences for following dynamical processes, although the physical description has the potential to discourage young scientists from delving into it. Here, the approach to the principles of Fluorescence Correlation Spectroscopy is based more on a phenomenological description of fluorescence fluctuations without ignoring recent theoretical and experimental findings.

Keywords fluorescence; confocal microscopy; autocorrelation

1. Introduction

Fluorescence Correlation Spectroscopy (FCS) is nowadays widely used for the determination of molecular parameters like diffusion coefficients or molecular masses, for identifying photochemical pathways and for quantifying concentrations of fluorescently tagged biomolecules [1-5]. This method is founded on the correlation analysis of fluorescence fluctuations which are observed when a small number of molecules or even single molecules are repeatedly excited in the detection volume of a confocal microscope. Other ways of analyzing these intensity variations exist, but they do not allow for assessing their duration [6-8]. However, typical time constants are required for distinguishing the different origins of these fluctuations. The most prominent sources are varying numbers of molecules in the detection volume due to diffusion and fluorescence intermittence due to quantum jumps of single molecules [9]. As the qualitative description is similar, I will focus on diffusion.

The theoretical treatment of FCS of diffusing particles is fully understood since the seminal work of Elson and Magde [10-12]. There, fluorescence fluctuations are traced back to concentration fluctuations in the excitation volume of a laser. Mathematical expressions of the concentration correlation functions are obtained by solving Fick's second law in cylindrical coordinates with the z-axis as the laser propagation direction. The detection probability is not uniform due to the Gaussian intensity shape of laser beams, and the fluorescence fluctuations finally are derived by convolving the intensity profile with the concentration correlation function. While the original work was focused on the two-dimensional diffusion perpendicular to the beam direction, recent experimental approaches are all based on the three-dimensional confinement of the detection volume in confocal microscopes. The third dimension imposes to approximate the axial intensity profile by a Gaussian on the theoretical treatment; similarly, the shape of the pinhole, which is used to suppress out-of-focus light, affects the resulting correlation function [13]. The widespread availability of confocal microscopes has made recording FCS traces easy. However, ignoring the mentioned physics might provoke misinterpretation of FCS data which otherwise can be overcome by careful performance.

In the following, I will describe how an autocorrelation function can be explicitly derived from fluorescence fluctuation and, shortly, how this is technically realized by hardware correlators. Artefacts due to the equipment are explained and also, how these can be overcome. The fitting process and

[1] E-mail: g.jung@mx.uni-saarland.de, phone: +49-681-302-64848

procedures to be undertaken for ensuring interpretability of the data are discussed before recent findings are presented as an outlook.

2. Fluorescence fluctuations

Fig. 1a sketches the situation of freely diffusing particles in the detection volume of a confocal microscope. Both the excitation probability and the detection efficiency peak in the centre where the focus of the laser and the image of the detection pinhole coincide. Two different trajectories of individual molecules are depicted, of which trajectory I traverses the volume of highest excitation and detection probability. Molecules following trajectory II only strike the detection volume. If we consider fluorescent molecules then trajectory I would correspond to a high fluorescence signal and trajectory II would lead only to a weak fluorescence signal.

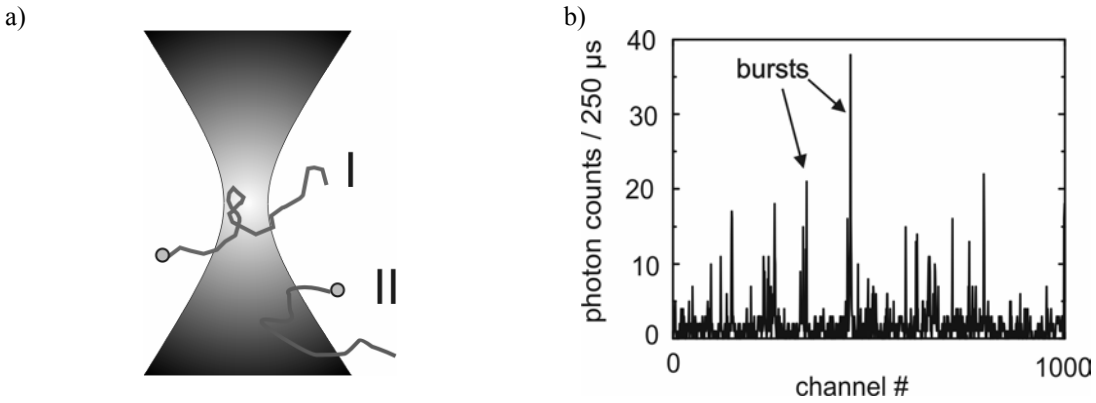

Fig. 1: Fluorescence fluctuations due to freely diffusing molecules through the focus of a laser beam. a) shows two explicit trajectories out of an infinite number of possibilities. The lighter the colour the higher are the detection and excitation probability. The propagation direction of the laser is vertical. b) shows the signal fluctuations due to diffusion. Molecules crossing the focus in a) are repeatedly excited while they are in the focus. The high intensity therefore causes numerous fluorescence photons within a short time interval, i.e. a photon burst. The channel width $\Delta t_{\#}$ in the time trace is 250 µs, the channel number is denoted by #.

The most frequently used detectors in FCS are avalanche-photodiodes (APD) which exhibit high detection efficiencies for fluorescence photons. Commercial single photon-counting APD modules generate electronic pulses of which each corresponds to one detected photon. Strong fluorescence signals therefore correspond to a large number of detection events and thus to a large number of photon counts within the channel dwell time $\Delta t_{\#}$ (fig. 1b). The mean residence time τ_{diff} of molecules in the detection volume is normally small and consequently the duration of the strong fluorescence signal is limited. The passage of a single molecule through the detection volume hence is detected by a photon burst.

The probability that a molecule only touches the detection volume in its outer region is higher than for a passage through the centre. This plausible explanation for the fact that large photon bursts are less often observed than weaker ones is hardly treatable by mathematics. There is an infinite number of possible pathways as diffusion is a statistical process, and the movement of individual species cannot be derived just from the fluctuation. Surfaces with equal detection probability are used to compute the photon signals, and these surfaces scale with the square of the distance to the centre of the detection volume [6]. By this, the distribution of photon counts within a fixed dwell time can be calculated. These so-called photon counting histograms are also analyzed with respect to the composition of mixtures of fluorescent species (fig. 2) [7, 8].

Considering a highly diluted sample, we can neglect interaction of individual particles, and Poisson statistics underlie the altering numbers of molecules in the detection volume. Absolute signal

fluctuations therefore are proportional to $\sqrt{N}$ with N as the average number of molecules in the detection volume. Thus, signal fluctuations relative to the mean fluorescence signal are proportional to $\sqrt{N}/N = 1/\sqrt{N}$. When we now calculate the variance g_0 of the relative fluctuations according to Poisson statistics, then g_0 scales with $1/N$. This indirect proportionality of g_0 to N is the reason why experimental techniques based on the analysis of fluorescence fluctuations are assigned to the single molecule techniques although the behaviour of individual species cannot be determined. On the one hand, all mentioned techniques average over a large number of single molecule events blurring the behaviour of outliers; on the other hand, this is a must for the correlation analysis being valid for fixing the physical quantities mentioned in the introduction.

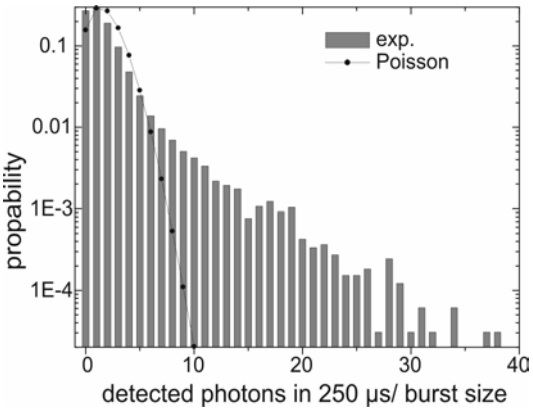

Fig. 2: Photon counting histogram from a longer time trace than depicted in fig. 1b. Small burst sizes are more likely than events with large count numbers. The probability for high count numbers within a channel however is higher than expected for pure photon noise with the same average count rate. Photon noise obeys a Poisson distribution (point dotted). These histograms can be analyzed for the concentration and composition of the solution [6-8].

Intensity fluctuations as depicted in fig. 1b might have different origins like light driven intersystem crossing [9], cis-trans isomerization [14], electron-transfer reactions or diffusion in and out of the detection volume. They might occur on different time scales and therefore are characterized by the corresponding time constants. Thus, freely moving fluorophors can be distinguished from those bound to membranes, antibodies, receptors and so on.

3. Correlation analysis

3.1 Manual correlation analysis of time traces

When the emission of an immobilized single molecule is observed in a long time trace then characteristic interruptions of the fluorescence emission are observed. This phenomenon of digital on-off transitions is called blinking and is a characteristic of single molecule emission [15]. The length of dark periods can be analyzed by a histogram analysis, whereas the length of the *on* periods reflects the probability for such a transition. By that, the rate constants for individual molecules are accessible. This kind of analysis however fails for most molecules at ambient conditions as photodestruction or accidental motion terminate the time trace before good enough statistics are available. Averaging over many molecules therefore is the only opportunity to determine the time constants of the processes which are responsible for the fluorescence fluctuations. Most important here is correlating the fluctuations. Even the analysis of the emission of single molecules benefits from it as this mathematical treatment eliminates the inherent noisy character of photon emission to a large extent.

Fig. 3a shows a cutout of a time trace like in fig. 1b. For an understanding how correlation function are generated by hardware or software correlators, we first calculate the average value of the fluorescence signal $S(t)$ over the whole time trace, $\langle S(t) \rangle$. The length of the time trace is denoted as T.

$$\langle S(t)\rangle = \frac{1}{T}\sum_{\#} S(t)\cdot \Delta t_{\#} \tag{1}$$

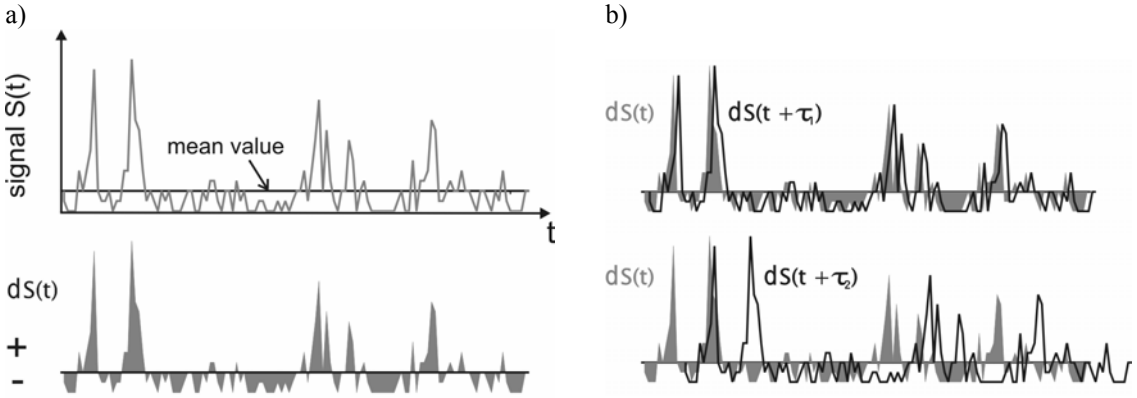

Fig. 3: The way from fluorescence signal fluctuations *dS(t)* in a) to the fluctuations, shifted by τ, *dS(t+*τ*)* in b). From the signal *S(t)* (grey line in a), *dS(t)* (grey, filled curves) can be calculated by subtracting the mean value <*S(t)*>. Shifting these fluctuation by a small increment on the time axis (τ_1) yields a curve (black line) in which the sign of *dS(t)* at each data point is similar to the original one. Longer shifts e.g. by τ_2 cancel this correlation.

The mean value <*S(t)*> is used to quantify the deviation of *S(t)* from it, i.e. the signal fluctuations *dS(t)*:

$$dS(t) = S(t) - \langle S(t)\rangle \tag{2}$$

Positive values of *dS(t)* can be interpreted as a fluorophore staying in the detection volume, negative values of *dS(t)* indicate the lack of fluorescent molecules in the centre of the detection volume.

Let us shift now *dS(t)* by a short increment of time, i.e. τ_1 in fig. 3b. The resulting curve *dS(t +* τ_1*)* exhibits the same sign of *dS(t)* at most datapoints. This finding is reasonable as it implies an interpretation via a conditional probability: if a molecule was in the detection volume at a certain moment, then the same molecule is highly likely to be observed shortly afterwards. It is also true that if molecules were absent then it is unlikely to register a strong signal in the next moment.

Now consider the curve which is shifted by τ_2. Here, positive fluctuations of the original curve (grey) coincide with positive or negative values of *dS(t +* τ_2*)* in a completely arbitrary manner. The same holds for negative deviations from the mean value. The statistical nature of the underlying processes (diffusion, blinking) guarantees that the fluctuations are "uncorrelated" on longer time scale than the typical period of these processes. The mean value <*S(t)*> is constant over time.

Mathematically, the correlation and its time dependent loss are analyzed by multiplying *dS(t)*•*dS(t +* τ*)* at each data point. Fluctuations with equal sign give positive products, whereas opposite signs generate negative values. Summing up these products over *T*, the duration of the experiment, therefore yields a positive sum for small τ, but decays to 0 for large τ (equation (3)). The unpredictability of the future fluorescence signal makes a positive or negative deviation from the mean value equally probable for large time lags.

$$G(\tau) \propto \frac{1}{T}\sum dS(t)\cdot dS(t+\tau)\cdot \Delta t_{\#} \tag{3}$$

Please note that the first data points of *dS(t)* as well as the last data points of *dS(t +* τ*)* must be ignored as there is no counterpart in the other curve. In principle, this should be accounted for by dividing the sum by $T\text{-}2\tau$. In most cases, however, is $T >> 2\tau$.

We are interested in the relative fluctuations and their typical decay time constant (cf. section 2), and therefore, each of the absolute signal fluctuations is divided by the signal's mean value, $\langle S(t) \rangle$.

$$G(\tau) = \frac{\langle dS(t) \cdot dS(t+\tau) \rangle}{\langle S(t) \rangle^2} \tag{4}$$

The resulting function G is a function of τ, $G(\tau)$. When fluctuations of a sample in the thermodynamic equilibrium are investigated, then the starting point of the experiment should not affect $G(\tau)$ and then any dependence of G on the "lab time" t is removed. In most application, however, this condition is not strictly fulfilled, but FCS can also be applied when the relaxation of the sample to the equilibrium e.g. by chemical reactions, adsorption, etc is distinctly slower than given by T or, at least, the typical time constant of the process under investigation.

3.2 Practical performance

Software or hardware devices which are used to measure correlation functions display these on-line during data recording. The described procedure for generating correlation functions "by hand" reveals two drawbacks of such a practical implementation.

Firstly, the definition of the fluctuations $dS(t)$ around $\langle S(t) \rangle$ requires that the latter value is known. This, of course, is only known *a posteriori*. A loop-hole is provided by multiplying the actual signals, $S(t)$, with $S(t + \tau)$

$$\begin{aligned} S(t) \cdot S(t+\tau) &= \left(dS(t) + \langle S(t) \rangle\right) \cdot \left(dS(t+\tau) + \langle S(t) \rangle\right) = \\ &= dS(t) \cdot dS(t+\tau) + dS(t) \cdot \langle S(t) \rangle + dS(t+\tau) \cdot \langle S(t) \rangle + \langle S(t) \rangle^2 \end{aligned} \tag{5}$$

The second and third term each are cancelled by integration over T.

Secondly, the time resolution is limited by the channel width of the time trace. Faster processes cannot be resolved. Reducing the channel width $\Delta t_{\#}$ is not an appropriate solution, as the capacity, which is necessary for data storage and multiplication in an increasing number of channels, easily exceeds the power of personal computers. Additionally, it is not mandatory or useful to analyse processes on a millisecond time scale with nanosecond time resolution!

Hardware correlators consist of an array of time channels, the width $\Delta t_{\#}$ of which increases with τ. While microsecond dynamics are measured with nanosecond channel widths, processes on the time scale of seconds are resolved with millisecond resolution. Thus, a manageable number of channels are sufficient to cover different dynamics spanning several orders of time dimension. This structure is represented in the logarithmic time axis of FCS curves (see below in fig. 4). In addition, the raw data are not stored permanently in the channels but continuously shifted in the electronic devices. The individual channels in these so-called FIFO memories (first-in-first-out) are read out for the multiplication at different τ as indicated in eq. (3) before the raw data get lost after the last and latest channel. Only the products and the total count number are summed up. The total count number is needed for an on-line averaging of $S(t)$.

The actual correlation curve which is displayed on-line on the computer screen therefore corresponds to

$$g(\tau) = \frac{\langle dS(t) \cdot dS(t+\tau) \rangle + \langle S(t) \rangle^2}{\langle S(t) \rangle^2} = 1 + G(\tau) \tag{6}$$

This explains why correlation curves decay to 1 and not to 0 (fig. 4). However, often 1 is subtracted from $g(\tau)$ or ignored. Even so, a commonly accepted rule for using minuscule or capital letters for $G(t)$ does not exist.

One result of the (quasi-)logarithmic width of the individual channels and consequently of the data points in fig. 4 is that the noise is larger at small τ than at large τ [16]. The probability to record photons scales with $\Delta t_{\#}$. Therefore data points on the millisecond time scale are built by roughly a factor of 10^6

more photons than data points on the nanosecond time scale. This explains why time and effort to determine diffusion coefficients (on a millisecond time scale) are smaller than for quantifying photophysical processes (on a nanosecond time scale).

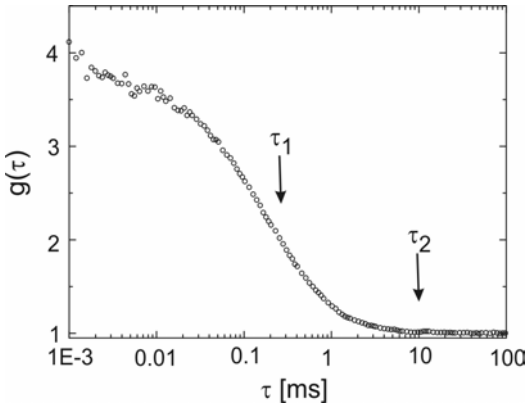

Fig. 4: Correlation of fluorescence fluctuations in the sample which was also used in the previous figures. g_0 is obtained by extrapolating $g(\tau)$ to 0 with subsequent subtraction of 1. The average number N of molecules in the detection volume is ~ 0.4. Small increments of τ only reduce g slightly (τ_1) compared to large τ. Please note that the time trace in fig. 1b is obviously not well suited to resolve the diffusional time constants as $G(\tau) = g(\tau)-1$ has already decayed to ½ of g_0. τ_1 and τ_2 approximate τ_1 and τ_2 in fig. 1b.

4. Determination of physical quantities

4.1 Diffusion as an example

The analysis of the correlation curves is similar for the different processes. In diffusion, two quantities can initially be extracted by fitting a formula to $g(\tau)$, i.e. the reciprocal value of the number of molecules in the detection volume, $1/N$, and the time constant of the fluctuations, τ_{diff}. As described in the derivation of eq. (3), negative and positive fluctuations around $\langle S(t) \rangle$ contribute to $G(\tau)$. It is therefore understandable that the typical time constant for fluctuations is composed of both time constants. However, this alone is not of relevance in the treatment of diffusion. The mechanisms which lead to balancing concentration gradients in a macroscopic experiment are the same like those of the microscopic fluctuations in equilibrium, i.e. the statistical nature of locomotion. Diffusion in and out of the detection volume are both described by Fick's second law for concentration gradients, and the outcome of solving the differential equation is the same that applies to the relaxation of the non-equilibrium condition. Although an exponential decay of $G(\tau)$ as a first guess is expected with a uniform detection profile, one has to remember that the detection probability of a molecule is not homogeneous across and along the detection volume. In their original work, Elson and Magde derived a hyperbolic decay of $G(\tau)$ by regarding the Gaussian intensity shape of the excitation laser beam [10].

$$G(\tau) = \frac{1}{N} \cdot \frac{1}{\left(1 + \frac{\tau}{\tau_{diff}}\right)} \tag{7}$$

The irradiation profile is characterized by the beam waist ω, which defines the decay time constant τ_{diff} together with the diffusion constant D.

$$\tau_{diff} = \frac{\omega^2}{4D} \tag{8}$$

Equation (7) is applied to all cases where fluctuations are mainly due to motions perpendicular to the light propagation direction. This also happens in confocal microscopes with low axial resolution, i.e. large detection pinhole diameters and distinct underfilling of the objective's backside aperture by the laser beam. The situation becomes more complicated when the axial resolution increases by expanding the excitation beam in front of the focussing objective lens and/or reducing the pinhole diameter. Then, fluctuations are caused by diffusion in and out in all three dimensions. This is tackled by Gaussian

approximations for the rectangular cross section of a pinhole aperture, which is imaged to the focal plane, and, more crudely, for the Lorentzian intensity profile along the laser propagation direction, but otherwise analytical solutions cannot be obtained [13]. Following the original work [10], eq. (9) is derived, in which ω_0 (instead of ω) contributing to τ_{diff} (eq. (8)), is now related to the optical resolution.

$$G(\tau) = \frac{1}{N} \cdot \frac{1}{\left(1 + \frac{\tau}{\tau_{diff}}\right)} \cdot \frac{1}{\sqrt{\left(1 + \frac{\tau}{C^2 \tau_{diff}}\right)}} \tag{9}$$

In eq. (9), the cylindrical shape of the detection volume is factored by C, the ratio of the vertical to the horizontal resolution. Fig. 5 compares fitting of eq. (7) and eq. (9) to the experimental data in fig.4.

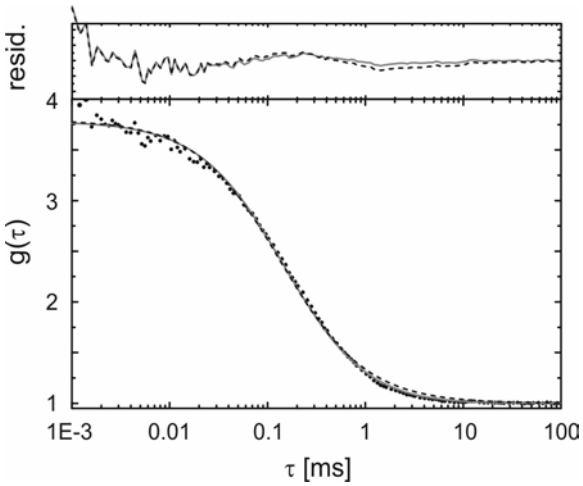

Fig. 5: Fitting of eq. (7) (dotted line; two-dimensional model) and eq. (9) (grey, full line; three-dimensional model) to experimental data. Use of eq. (9) slightly improves the fit. In consequence, the detection volume can be determined to 1.5 femtoliters.
Often, fit residuals, i.e. the deviation of the experimental data from the fit, are plotted for judging the quality of the fit. Irregularities in the residuals indicate some mismatch of model and data. At least in parts, this mismatch can be traced back to the approximations made for deriving eq. (9). Simulations show that significant differences between 2D and 3D-fits only appear at C-values below 4 - 5, i.e. in the limit of the highest achievable axial resolution

The advantage of applying eq. (9) is that a volume can be defined in which N and consequently the concentration can be determined.

4.2 Artefacts and Caveat

Additional features in correlation curves might arise from instable lasers, the power supply system, etc, but also from intrinsic properties of the used detectors [17]. Afterpulsing and dead-time effects of these can be overcome by using two detectors with equal detection probability [18]. Subsequent crosscorrelation of the output of both modules, i.e. $dS(t)$ and $dS(t+\tau)$ are stemming from different detectors, contains the same information like the autocorrelation which has been discussed so far.
More severe and harder to recognize are curve distortions due to imperfection of the optical system. Refractive index mismatch, varying cover glass thickness and intensity dependent saturation phenomena can alter τ_{diff} due to enlargement of ω_0 in eq. (8) or even change the shape of the decay of $G(\tau)$ [19]. Photobleaching has similar effects.

Even when special care is applied to remove these influences while recording FCS traces, the quantities N and D returned by a fit can deviate from estimates. The diffusion constant D depends on the viscosity in the Einstein-Stokes-relation which itself is strongly temperature-dependent. For the determination of a realistic N, one has to correct for background photons which can originate from Raman or fluorescent impurities [13, 18]. Furthermore, examination of eq. (3) shows that $G(\tau)$ is composed of the square of fluorescence intensities. This nonlinearity in $G(\tau)$ favours strongly fluorescent molecules over weakly fluorescent species [5]; therefore, non-stoichiometric labelling intensifies the nonlinearity.

In summary, FCS is a valuable spectroscopic method for assessing molecular parameters, but the ease of performance might entail some hazards of misinterpretion.

Acknowledgements The support by the German Science Foundation (DFG) is gratefully acknowledged. I would like to thank Ms. Michaela Jacob for careful reading and criticism.

References

[1] D. Grunwald, M. Cardoso, H. Leonhardt, V. Buschmann, Curr. Pharm. Biotechnol. **6**, 381 (2005).
[2] M. Gosch, R. Rigler, Adv. Drug Deliv. Rev., **57,** 169 (2005).
[3] O. Krichevsky, G. Bonnet, Rep. Prog. Phys., **65**, 251 (2002).
[4] P. Schwille, Cell Biochem. Biophys., **34**, 383 (2001).
[5] M. Eigen, R. Rigler, Proc. Natl. Acad. Sci. USA, **91,** 5740 (1994).
[6] J. Fries, L. Brand, C. Eggeling, M. Köllner, C. Seidel, J. Phys. Chem. A **102** 6601 (1998).
[7] Y. Chen, J. Müller, P. So, E. Gratton, Biophys. J.**77**, 553 (1999).
[8] P. Kask, K. Palo, D. Ullmann, K. Gall, Proc. Natl. Acad. Sci. USA **96**, 13756 (1999).
[9] Th. Basché, S. Kummer, Ch. Bräuchle, Nature **373**, 132 (1995).
[10] E. Elson, D. Magde, Biopolymers **13**, 1 (1974).
[11] D. Magde, E. Elson, W. Webb, Biopolymers **13**, 28 (1974).
[12] D. Magde, E. Elson, W. Webb, Phys. Rev. Lett. **29**, 705 (1972).
[13] R. Rigler, Ü. Mets, J. Widengren, P. Kask, Eur. Biophys. J. **22**, 169 (1993).
[14] J. Widengren, P. Schwille, J. Phys. Chem. A **104**, 6416 (2000).
[15] G. Jung, E. Kneuper, L. Seyfarth, A. Zumbusch, in: Science Technology and Education of Microscopy: an overview (ed. by A. Méndez-Vilas), Formatex, Spain, 2003, pp. 297-303.
[16] T. Wohland, R. Rigler, H. Vogel, Biophys. J. **80** 2987 (2000).
[17] J. Rarity, T. Wall, K. Ridley, P. Owens, P. Tapster, Appl. Opt. **39**, 6746 (2000).
[18] J. Widengren, Ü. Mets, R. Rigler, J. Phys. Chem. **99**, 13368 (1995).
[19] J. Enderlein, I. Gregor, D. Patra, T. Dertinger, U. Kaupp, ChemPhysChem **6**, 2324 (2005).