Single-Phase Single-Molecule Fluorescence Correlation Spectroscopy (SPSM-FCS)

Zeno Földes-Papp  
Medical University of Graz, Graz, Austria

Gerd Baumann  
Department of Mathematical Physics, University of Ulm, Germany and  
Department of Mathematics, Faculty of Basic Sciences,  
The German University in Cairo, Egypt

Masataka Kinjo  
Mamoru Tamura  
Laboratory of Supramolecular Biophysics, Research Institute for Electronic Science,  
Hokkaido University, Sapporo, Japan

INTRODUCTION

Today, we are largely living in Albert Einstein’s macro and micro universe. And it all goes back to four groundbreaking papers published by Einstein in 1905 in the *Annalen der Physik Leipzig*—three major works and an addendum. In June 1905, he explained how light works, in July of that year he explained how matter works, in September he explained how space and time obey the speed of light, and in November 1905 he submitted his $E = mc^2$ paper, building on his July results. Albert Einstein’s theories have all proven to be true and essential to everyday applications.

In the context of medical genomics and proteomics, Einstein’s July 1905 ideas are alive, but they are better known than understood. He proposed an atomic explanation for the unexplainable random motions of particles such as DNA or protein molecules suspended in liquid. Because diagnostic assays in medical genomics and proteomics are now extremely sensitive, there can be a need to quantify molecule numbers below 10.

Fluorescence correlation spectroscopy and two-color fluorescence cross-correlation spectroscopy are used to perform the measurements because both of these experimental methods directly count the molecule number. For example, the measured amplitude (minus 1) of the normalized fluorescence intensity correlation curve of 0.1 corresponds to an assemble average of $N = 10$ molecules per confocal probe region, a signal amplitude of 0.2 to 5 molecules, and a signal amplitude of 1 to 1 molecule.

TECHNICAL DESCRIPTION

There is the need for algorithms to estimate the detection of just one (“self-same”) fluorescent molecule in single phases such as solutions or membranes. The requirements for such algorithms are very stringent:

1. The estimation of the probability that just one, the self-same, molecule is measured during the collection time (measurement time) must be based on the data collected from one particular measurement only. It cannot be based on a set of related measurements because there are variations in the probability from one measurement to another.
2. The estimation must be capable of detecting the self-same molecule without immobilization or hydrodynamic flow.
3. The estimation must be computed from the measurement of the molecule numbers $N$, and not from the emitted fluorescent photon bursts themselves.

Requirement 3 may appear simple, but it is important and of outstanding practical value because the most common measurements do not yield emitted photon bursts or lifetime statistics. A convenient way to meet requirement 3 is to use fluorescence correlation spectroscopy and two-color fluorescence cross-correlation spectroscopy.1–9 Single (solution)-phase single-molecule fluorescence auto- and two-color cross-correlation spectroscopy (SPSM-FCS) is based on the detection of the Brownian motion of fluorescent molecules and Poisson distribution analysis depending on the average frequency, $C$, as well as the presence in...
The very tiny detection volume of a single fluorescent molecule, on the molar concentration, $c_m$, of the bulk phase (e.g., solution), and on a specified molar bulk concentration, $c_m^*$, below which a probabilistic assignment is meaningful.

**THE SELF-SAME MOLECULE LIKELIHOOD ESTIMATORS**

**The General Case**

We have previously proposed the general case of the selfsame molecule likelihood estimator, i.e., the Eq. (11) in Ref.[1], which fits the three requirements:

$$P\left(\bigcap_{i=1}^{2} A_i\right) = 1 - \prod_{i=1}^{2} P(A_i) = 1 - P_1 \cdot P(q\leq q<+\infty)$$  \hspace{1cm} (1)

This selfsame molecule likelihood estimator $P\left(\bigcap_{i=1}^{2} A_i\right)$ gives the probability for the independent events of finding (arrival of) a single fluorescent molecule ($A_1$) in the confocal detection/observation volume (confocal probe region) and of moving (departure) of this single fluorescent molecule ($A_2$) over a fixed distance (lower limit $\bar{z}$). $P_1$ is the measured probability of finding a single fluorescent molecule in the detection volume and $P(q\leq q<+\infty)$ is the frequency distribution of the diffusive spreading for the same single molecule. The measured dimensions of the detection volume, e.g., the focus of the excitation laser beams, give the lower limit of distance $\bar{z}$. There is in fact an entire family of valid likelihood estimators of the same structure as Eq. (1), although they may differ with respect to the properties of $P(q\leq q<+\infty)$. Finding the most suitable estimator to use is still an interesting field of investigation.

From Eq. (1), it is evident that the probability distribution of finding a single fluorescent molecule in the probe region is of paramount importance. The random molecule fluctuations make it possible to assign a probabilistic value to such a property in the detection volume of a bulk phase below a specified (“critical”) molar concentration $c_m^*$ of that bulk phase. Our analysis is straightforward and profound.[2,3] We say that if, in the absence of a macroscopic concentration gradient, the specific fluorescent molecules tend to a stable uniform steady state in the bulk phase, then the numbers of the molecules (events) are Poisson-distributed in the unit volume of detection as follows. We record the absolute number of events $X$ occurring in an interval of time $T$ (infinite number of periods of $T$ units of time). The events ($\chi$) happen with a constant detection probability $\Pi$ per unit time. $\Pi$ is related to molecular properties of the fluorescent molecule and to instrumental parameters of the measuring device. Finally, we represent the SPSM-FCS experiment by a probability (“observable”) that $\chi$ events take place in the interval $T$:

$$P(X \leq \chi; \Pi \cdot T = C)$$

$$= \int_0^T \Pi \cdot e^{-\Pi t} \cdot \frac{1}{(\chi - 1)!} \cdot \Pi^{\chi - 1} \cdot (T - t_0)^{\chi - 1} \cdot e^{-\Pi T} \, dt_0 = \left(\frac{(\Pi \cdot T)^\chi}{\chi!}\right) \cdot e^{-\Pi T}$$  \hspace{1cm} (2)

$\Pi \cdot T$ is the average number of specific fluorescent molecules in the unit detection volume $\Delta V$, meaning the mean value of $\chi$. The value $\chi$ is always a small integer, e.g., $\chi = 1$ in Fig. 1. Now, we use $C$ to denote the (true) mean value of the population of specific fluorescent molecules (the average molecule number) in the confocal femtoliter and less detection volume and summarize our approach by[1-3]

$$P(X \leq \chi; C) = P_\chi = \frac{C^\chi \cdot e^{-C}}{\chi!}$$  \hspace{1cm} (3)

We have so established this novel, condensed, time-averaged Poisson-distributed relationship, represented by the above equation.

The average molar concentration of other molecules of the same kind in the bulk phase is $c_m \equiv C/(N_A \cdot \Delta V) \approx N/(N_A \cdot \Delta V)$ under the experimental conditions $C \ll e^{-C}$ with $N_A$ as Avogadro’s number of $6.023 \times 10^{23}$ mol$^{-1}$, and $\Delta V$ the detection volume.[2] Hence, the probability $P_1$ that the detection volume contains a single (specific) fluorescent molecule, $\chi = 1$, is according to Eq. (3).

$$\ln \{N \equiv P(X = 1, \Pi \cdot T = C)\} = \ln \{N = P_1\} = \ln C - C \quad \text{(criterion 1)}$$  \hspace{1cm} (4)

and the probability $P_2$ that the detection volume contains two (specific) fluorescent molecule, $\chi = 2$, reads

$$\ln \{P(X = 2, \Pi \cdot T = C)\} = \ln \{P_2\} = 2 \cdot \ln C - \ln 2 - C \quad \text{(criterion 2)}$$  \hspace{1cm} (5)

The second criterion [Eq. (5)] now describes the “analytical sensitivity” that the detection volume contains “a single fluorescent molecule.” As previously shown in Table 1 of Ref.[3], it was experimentally confirmed that the Poisson probability $P_1$ for the measured occurrence
of a single-molecule event is $P_t = N \cong C$ under the experimental conditions $C \ll e^{-C}$.

Taken together, the main difference to other Poisson analyses in the literature, e.g., Ref. [10], is that the final expressions [Eqs. (3)–(5)] no longer contain the detection probability $\Pi$. It cancelled out. The use of the expressions for single-molecule experiments depends on a specified molar bulk concentration $c_m$ (“critical” concentration) at which the probabilistic assignment is meaningful. $c_m$ is about 1 nM for a detection volume of about 0.2 fl ($0.2 \times 10^{-15}$ L). With the equation [Eq. (1)], we have developed an estimator that takes symmetries of the probe region explicitly into account.

THE SPECIAL CASE

The self-same molecule likelihood estimator [Eq. (1)] is considered for an instantaneous line source with radial diffusion in an axially symmetric cylindrical coordinate system of the three-dimensional confocal probe region. In practical applications, it is easier and faster to use the approximation of the symmetry of the confocal detection volume. In this case, the estimator is of a straightforward analytical form [Eq. (10a)] that can be easily used by an experimenter who is not familiar with the physics and mathematics behind the formulas.

In three dimensions, Cartesian coordinates are the rectangular coordinates $(x, y, z)$. An alternative is cylindrical polars, which involve $q, \phi$, and $z$, as shown in Fig. 2. We note that $q$ and $\phi$ exactly correspond to the polar coordinates in the two-dimensional polar case to describe points in the $x$-$y$ plane, with $z$ being simply added. Therefore, the $z$-axis will also be used to describe the diffusion along a membrane by the same $q$ and $\phi$ in two-dimensional planes (sections) of this membrane. The relationships between the two sets of coordinates are

$$
\begin{aligned}
x &= q \cdot \cos \phi \\
y &= q \cdot \sin \phi \\
z &= z
\end{aligned}
$$

where $q$ is the distance of the projection of the three-dimensional space point into the $x$-$y$ plane, rather than that from the origin as in spherical polars. Which coordinate system is more appropriate depends on the nature of the problem and its symmetries. The three-dimensional confocal detection volume is schematically depicted in Fig. 1. In good approximation, it is cylindrically symmetric rather than spherically symmetric. Therefore, we are motivated to use cylindrical polars. We further assume isotropic media in good approximation, where diffusion properties are not different in different directions.

In terms of cylindrical polars $q, \phi$, and $z$ for an volume element of a cylinder with sides $dq, d\phi$ and $dz$ and radial diffusion in space (Fig. 1), the diffusion equation for a single molecule becomes

$$
\frac{\partial n}{\partial t} = D \cdot \left( \frac{\partial^2 n}{\partial q^2} + \frac{1}{q} \cdot \frac{\partial n}{\partial q} \right)
$$

where $n$ is the density function of a single fluorescent molecule, $D$ is simply the diffusivity or the diffusion coefficient, and $q = \sqrt{x^2 + y^2}$ is measured from its...
Single-Phase Single-Molecule Fluorescence Correlation Spectroscopy (SPSM-FCS)

Fig. 2 The single molecule $P$ in three dimensions has cylindrically polar coordinates $(q, \phi, z)$. $Q$ is the foot of the perpendicular from the single molecule $P$ into the focal $x$–$y$ plane of the detection/observation volume. $\phi$ is the angle between $0Q$ and the $x$-axis. $q$ is the length of $0Q$.

The single molecule likelihood estimators are then a function of the radius $q$ and the time $t$ only:\[1,4\]

$$n(q, t) = \frac{1}{4\pi Dt} \cdot \exp\left\{-\frac{q^2}{4Dt}\right\}$$

(8)

Hence, the probability $P(\xi \leq q < +\infty)$ that a fluorescent molecule is outside a boundary $q = \xi(t)$ at time $t$, meaning outside the lower limit of distance $\xi$, is given by integration over the radial distance $q$\[1,4\]

$$P(\xi \leq q < +\infty) = P(\xi) = \frac{1}{2\pi} \cdot \exp\left\{-\frac{\xi^2}{4Dt}\right\}.$$  

(9)

With equation [Eq. (9)], the special case of the novel self-same molecule likelihood estimators $P(\bigcap_{i=1}^{N} A_i)$ is\[1,4\]

$$P\left(\bigcap_{i=1}^{N} A_i\right) = 1 - P_1 \cdot P(\xi \leq q < +\infty)$$

$$= 1 - \frac{N}{2\pi} \cdot \exp\left\{-\frac{\xi^2}{4Dt}\right\} \quad \text{(criterion 3)}$$  

(10a)

Here $N$ is criterion 1 [Eq. (4)] and is measured, for example, by counting the molecule number per confocal probe region with fluorescence techniques such as fluorescence correlation spectroscopy. First criteria 1 and 2 must be fulfilled, and then criterion 3 can be used. All quantities on the right hand side of criterion 3 [Eq. (10a)] are obtained by the experiment, e.g., by SPSM-FCS.

**HOW TO USE THE SELF-SAME MOLECULE LIKELIHOOD ESTIMATOR**

The Poisson probability $N < 1$ with $P_1 = N \equiv 0.14$ of finding a single two-color molecule complex was observed in solution. The Poisson probability of finding two color molecule complexes is about 1% and can be neglected. The average diffusion coefficient of the molecule complex was experimentally determined with $D = (o_{x,y}^{\text{blue}} + o_{x,y}^{\text{red}})/8 \cdot \tau_{gr} = 4.5 \times 10^{-12} \text{m}^2/\text{sec.}$ The measurement times were kept at 60 sec. The novel quantity $P(A_1 \cap A_2)$ was determined to be $N_{\frac{0.159}{2\pi}} \cdot \exp\left\{-\frac{\xi^2}{4Dt}\right\} = 0.14\times \exp\left\{-\frac{(0.159 \times 10^{-6}[\text{m}^2]}{8(4.5 \times 10^{-12}[\text{m}^2/\text{sec}])}\right\} = 22.28 \times 10^{-3}$, where $\xi = o_{x,y}^{\text{blue}} = 0.159 \mu m$. Hence, the probability that the same large immune complex is measured under the given experimental reaction conditions, is $1 - 0.02228 = 0.97772$. Just one individual complex molecule of an autoantibody target in serum was assayed with a confidence of about 97.8%, but the value does not describe the meaningful time $T_m$ that one can study the self-same molecule; see legend to Fig. 1 and the Eq. (11) which is found/derived in Ref.\[13\] for the first time.

A change in the symmetry of the confocal detection volume, e.g., by a “cellular” distortion, would change the calculation. For an additional angular integration of $n(q, \phi, t)$ in the above example, the calculation result means that the probability within the angular range $0 \leq \phi \leq 2\pi$ (Fig. 2) is equally distributed, and therefore it is not subject to changes with a confidence of 86.0%. However, the measurements do not permit any angle resolution. The same reasoning holds true along the $z$-axis of the given confocal detection volume with the probability of about 99.9%, where $2 \cdot o_{z}/\text{unit} = 2 \cdot o_{z} = 1.91 \times 10^{-6}$ is a ratio and equals the double numerical value of the extracted $o_{z}$ at a given size of the confocal femtoliter and less probe region (Fig. 1). But in fact, the cylindrical polaris are more appropriate to consider the diffusive rate for a molecule that does not begin at the origin $(x, y, z) = 0$. If a molecule is near the upper or lower edge of the confocal probe region in the beginning of its diffusive spreading, i.e., $z \neq 0$ in Figs. 1 and 2, it is also...
taken into account. From the above special cases, the following property of the self-same molecule likelihood estimators $0 \leq P(\bigcap_{i=1}^{\infty} A_i) \leq 1$ is asserted.

$$P\left(\bigcap_{i=1}^{\infty} A_i\right) = 1 - P_1 \cdot P(\xi < q < +\infty)$$

$$= 1 - P_1 \cdot \int_{z_1}^{\infty} \int_{0}^{q_\xi} \int_{0}^{+\infty} n(q, \phi, z, t) \cdot q \cdot d\phi \cdot dz,$n(q, \phi, z, t)d\phi \cdot dz,$

$$= 1 - P_1 \cdot \int_{z_1}^{\infty} \int_{0}^{q_\xi} \int_{0}^{+\infty} n(q, \phi, z, t)dx \cdot dy \cdot dz.$$ (10b)

During measurement times $T$, reentries of a single molecule in the probe region are given rise to fluctuation phenomena, such as molecule number fluctuations at the single-molecule level. The concentration dependence of the meaningful time $T_m$ (Ref.[13]) that one can study bimolecular reactions of the self-same molecule in the probe region is first found in Ref.[13].

$$T_m = \frac{\tau_{diff}}{c_m N_A \Delta V \cdot \exp\left(-c_m N_A \Delta V\right)} \quad \text{(criterion 4)}$$ (11)

The criterion 4 specifies the time parameter $t$ in criterion 3 [Eq. (10a)]. The experiments performed so far in Ref.[2–8] and the theory are consistent with the assumptions made.

**SPECIFICITY**

The developed algorithms make it possible to distinguish between the observation of single molecules one by one and the observation of just one, the self-same, individual molecule at a time, without immobilization, photon count statistics, or hydrodynamic flow. The observation of an individual, complex biological molecule in its natural environment is also possible over extended periods of time.[13] Progress in genomic, immunogenomic, and proteomic testing first means sharpening the specificity.

**SENSITIVITY**

The results demonstrate that single-molecule analytical sensitivity is essential for accurate molecule number quantification at the “single-molecule level.” If single-molecule sensitivity is not achieved, an unambiguous and/or incorrect molecule number determination may result.[8] The higher the sensitivity of an assay, i.e., the lower the measured $N$ value with $N < 1$, the better the performance of the confocal fluctuation measurements. This can be achieved by reducing the size of the detection volume/probe region (see Eq. (11) and Ref.[13]). A 200-fold increase in sensitivity was obtained as compared to the conventional ELISAs on solid phase.[5]

**REPRODUCIBILITY**

With these developed self-same molecule likelihood estimators, as with other statistical analyses, a sufficient number of replicate measurements are required to achieve reasonably good accuracy. SPSM-FCS provides the experimental platform to perform high-throughput measurements and the sensitivity necessary for single-molecule detection and identification in medical genomics and proteomics.[3]

**LIMITS**

According to the Polya theorem,[11] the probability of ultimate return to the starting point for any individual molecule (lattice) is 1 in one dimension and two dimensions and less than 1 in three or more dimensions.[11,12] This is the generalization of ordinary random walks of “individual molecules” for studying the distribution of end points. The distribution of any coordinate of the end point of a walk approaches a Gaussian value asymptotically in all dimensions. In addition, the distributions of different Cartesian coordinates $(x, y, z)$, rectangular coordinates, of the end point are independent and the distribution of the end point in space is thus spherically symmetrical. Nevertheless, the analytical approach taken with the special case of the novel self-same molecule likelihood estimators is very useful if we do not proceed to the infinitive limit of collection time ($t \to \infty$), i.e., if we restrict ourselves to finite measurement times as demonstrated in this entry and the related original articles[1–9] (Földes-Papp, Personal communication). The smaller the measured experimental $N$ value with $N < 1$, the more likely is the condition for the self-same single-molecule regime.

**CLINICAL APPLICATIONS**

A number of instrumentation challenges, collateral developments in chemistry and biochemistry, and new strategies for the investigation of complex biological molecules are needed to carry out single-molecule studies in complex chemical and biological environments. The main outcome of the experiments summarized in the section titled “How to Use the Self-Same Molecule Likelihood Estimator” is that an individual
two-color molecule complex, consisting of green-tagged antigens, sandwiched autoantibodies, and red-tagged detecting antibodies, is observed. The meaningful time $T_{\text{meq}}$ (Eq. 11) describes the time that one can study the individual, self-same molecule. The detection is improved by two orders of magnitude over conventional state-of-the-art ELISAs. It allows analysis of smallest amounts of biomolecules that are difficult to handle experimentally. The results point toward functional dose–response curves on the expected heterogeneity of the reaction behavior of individual biomolecules in medical proteomics. To get the next self-same fluorescent molecule, the position of the confocal probe region within the sample is moved, for example, by the joystick.

Single-molecule fluorescence methods enable a new class of nucleic acid assays that are not possible with PCR-driven methods. The novel experimental strategy of disease gene identification with unamplified double-stranded genomic target DNA in solution has been shown for the first time. It dramatically improves statistical accuracy and accelerates genotyping with bilabeled nonlinear probe oligodeoxyribonucleotides (molecular beacon superquencher probes) in a homogeneous assay format. The data will enable a distinction at the single-molecule level between underlying haplotype pairs and, ultimately, between “high risk” and “low risk” alleles at femtomolar allele concentrations and less without amplification or transcription. Potentially any nucleic acid molecule is a target for study and clinical application. This work will encourage others to apply single-molecule fluorescence methods to genomic analyses.

Single (solution)-phase single-molecule fluorescence auto- and two-color cross-correlation spectroscopy can be used for proteomics, nucleomics, glycomics, and epigenomics, providing direct information about molecular interactions of freely diffusing target molecules, such as aptamer–protein interactions, protein–protein interactions, and peptide antibody mimetic (PAM)–protein interactions. Further, SPSM-FCS can be a solution of limitations of microarray-chip-based platforms. It offers a panel of tools for this research field in biology, medicine, and pharmacology. The self-same molecule likelihood estimators are an enhanced tool for the rapid finding of appropriate experimental conditions to perform measurements with just one molecule.

CONCLUSIONS

The theory is mainly based on two facts: i) the existence of a continuity equation and ii) the existence of a probability current. Probabilistic models are of great value in the description of single molecules but their ultimate justification rests on approximations and experimental studies. This new development may lead to a clarification of concepts used in widely different fields of medical genomics and proteomics such as biochemistry, molecular biology, and immunology. In all these fields, we find again and again the same questions: what is the meaning of single-molecule approaches? Are they just another tool? Can we be sure of measuring just one molecule, i.e., the self-same molecule?

ACKNOWLEDGMENT

This article was originally part of the lecture given by Zeno Földes-Papp at the symposium “Single-Molecule Detection and Spectroscopy: A 15-Year Perspective” at the PITTCON 2005 Conference, February 27-March 4, 2005 held in Orlando, Florida, U.S.A. Zeno Földes-Papp wants to thank Richard A. Keller, Los Alamos National Laboratory, USA, for discussions and helpful comments on this article.

REFERENCES


