Dealing with reduced data acquisition times in Fluorescence Correlation Spectroscopy (FCS) for High-Throughput Screening (HTS) applications

Lloyd M. Davis[a], David A. Ball[†a], and Peter E. Williams[†a],
Kerry M. Swift[b], and Edmund D. Matayoshi[b]

[a]*Center for Laser Applications, University of Tennessee Space Institute, Tullahoma, TN 37388*
[b]*Abbott Laboratories, 100 Abbott Park Road, Abbott Park, IL 60064-6114*

ABSTRACT

Fluorescence Correlation Spectroscopy (FCS) may be used to assay the binding of drug-like ligands to signaling proteins and other bio-particles. For High Throughput Screening (HTS), a competitive format is typically used in which the binding of an unlabeled compound results in displacement of a fluorescently labeled ligand. Unweighted curve-fitting of the normalized autocorrelation function (ACF) to a two-diffusion-component model can resolve the fractions of free and bound ligand if the diffusion rates differ sufficiently and if the experimentally estimated ACF has adequate precision. However, for HTS (and also for intracellular FCS studies) it is desirable to minimize the experimental data collection time. In this case, the precision of the ACF is limited and it becomes important to account for the statistical features of the ACF estimate when designing an assay. The errors at different points in the estimated ACF are correlated and hence least-squares fitting methods are not statistically rigorous. We compare different methods for estimating and curve-fitting the ACF from the raw data of short duration FCS measurements. The methods are applied to data from experiments to assay binding of Alexa-488-labeled Bak peptide with Bcl-x_L, which is an intracellular protein that acts to protect against programmed cell death. We present results from a detailed Monte Carlo simulation of the experiment, which is useful for validating short-duration assay capabilities. We also discuss the measurement of changes in steady state fluorescence anisotropy due to restricted rotational diffusion upon binding, which provides a complementary assay.

1. USE OF FLUORESCENCE POLARIZATION AND FLUORESCENCE CORRELATION SPECTROSCOPY FOR DRUG-BINDING ASSAYS IN HIGH-THROUGHPUT SCREENING

The initial task in high-throughput pharmaceutical drug screening is to sort through hundreds of thousands of drug-like library compounds to find those that bind or interact with a target protein or bio-molecule, which has been identified from prior molecular-biology knowledge [1]. Miniaturization of sample volumes to conserve reagents and homogeneous solution assay are generally preferable in the initial screening measurements. The overall time for high throughput screening (HTS) may be reduced by multiplexing the initial measurements. For example, each sample may contain many different compounds, with each compound occurring in two or more different samples, and matching of the binding assay hits may be used to identify the particular compound responsible for each hit. Further measurements and research can then be performed on the identified compounds.

Figure 1 illustrates typical data from a HTS run. An experimental measure is made of the degree of binding to the target that occurs for any of the compounds within each sample. Systematic errors in the binding measure are oftentimes present. These may arise from many causes, including precipitation of solution components, uncertainties in the concentrations of solution components, unwanted interactions between test compounds, and unaccounted sources of signal such as autofluorescence [1]. Such errors account for apparent binding results that are less than 0%, or greater than 100%, as seen in Figure 1. Nevertheless, it is still possible to identify many compounds which consistently yield apparent binding, and which warrant further assays. However, these hits are usually only expected to lead synthetic chemists toward new and better compounds. The success of the drug discovery/design process depends on further medicinal chemistry and structural biology.

[†] DAB and PEW acknowledge travel support from the University of Tennessee Center in Chemical Physics

Microarrays and Combinatorial Technologies for Biomedical Applications: Design, Fabrication, and Analysis, Dan V. Nicolau, Ramesh Raghavachari, Editors, Proceedings of SPIE Vol. 4966 (2003) · 1605-7422/03/$15.00

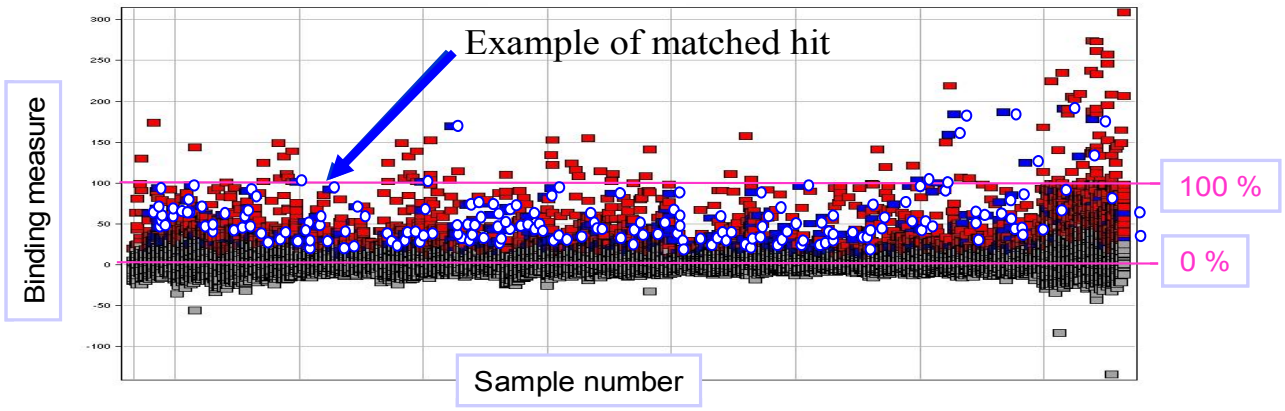


Figure 1: Typical data from a high-throughput screening run.

Fluorescence detection methods form the basis of many assays for determining whether interactions occur between the target biomolecule and any of the drug-like compounds within each sample. Most importantly, fluorescence provides exquisite sensitivity and requires only microliter sample volumes and nanomolar concentrations of components. In order to test hundreds of thousands of different drug-like compounds with the same assay, a competitive binding strategy is generally employed. First, a ligand which is known to bind to the target biomolecule must be fluorescently labeled. Optionally, the biomolecule itself may be fluorescently labeled using a distinct dye [1]. The labeled ligand is added to the target biomolecule solution at a concentration at which about 50% becomes bound. Under these conditions, the fraction of binding is most sensitive to competitive binding by other agents. Then, as illustrated in Figure 2, if a drug-like molecule that can also bind to the target is added, some of the labeled ligand becomes displaced into free solution.

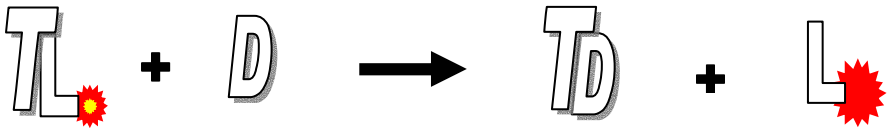


Figure 2: In a competitive binding assay, a drug-like molecule (D) is added to a solution containing fluorescently-labeled ligand (L) bound to a target biomolecule (T), thereby resulting in displacement of the labeled ligand, and a possible change in its fluorescence properties, or rotational or translation diffusion rates.

The fluorescence brightness, lifetime, or other properties may change, but more generally, the translational and/or rotational diffusion of the ligand become faster. In the latter case, Fluorescence Polarization (FP) may be used to monitor rotational diffusion and FCS may be used to monitor translational diffusion [2]. For Einstein-Stokes diffusion and approximately globular molecules, the translational diffusion scales inversely with the radius, whereas the rotational diffusion scales inversely with the radius cubed. Hence the rotational diffusion is a more sensitive indicator of binding, scaling inversely with molecular weight, whereas translational diffusion only scales inversely with the cube root of molecular weight. Thus to achieve a two-fold change in the translational diffusion, the target biomolecule must be at least 7 times the size of the ligand. On the other hand, FP can only monitor changes in the rotational diffusion provided the fluorescent label is rigidly bound to the ligand. It may be advantageous to measure the time resolved fluorescence anisotropy [3]. Oftentimes the label is bound to the ligand with partially restricted motion, but is free to rotate about a particular axis, such as around a covalent link. In such cases, the time resolved fluorescence anisotropy would exhibit multiple exponential decay components, whereas the time-averaged or "steady-state" anisotropy would be dominated by the faster component, which is not altered by the binding of the ligand to the biomolecule. Further, if FP is to successfully monitor changes in ligand-target binding, the fluorescence lifetime of the label must be longer than the slow-axis rotational lifetime of the ligand when it is bound to the target biomolecule. With the above considerations, FP and FCS are best considered as complementary rather than competing tools for assaying the competitive binding of drug-like compounds to a target [4]. Other complementary methods, which are particularly helpful when mass changes due to binding are not detectable by FP or FCS, include assays based upon counting photons within successive time intervals, which are applicable when binding results in at least a 2-fold change in fluorescence brightness, such as analysis of the photon counting histogram (PCH) [5], or fluorescence intensity distribution analysis (FIDA) [6].

2. EXAMPLE: Bcl-x_L

Because of its implication in a wide range of clinical disorders, including Cancer, autoimmune diseases, viral infections, and neuro-degenerative disorders, Bcl-x_L may be a useful signaling protein molecule to target for drug discovery efforts. Bcl-x_L is a protein from the Bcl-2 family that acts to protect against cellular apoptosis by complexing with pro-apoptotic Bcl-2 proteins, such as Bax, Bak, Bid, and Bad [7]. The latter stimulate mitochondrial release of cytochrome c, which activates caspases that in turn are the executioners of programmed cell death. Researchers at Abbott Laboratories have labeled a synthetic 16-residue Bak peptide with the fluorophore Alexa 488 from Molecular Probes, as shown in Figure 3, and have determined the structure of the Bcl-x_L + Bak-ligand complex [8].

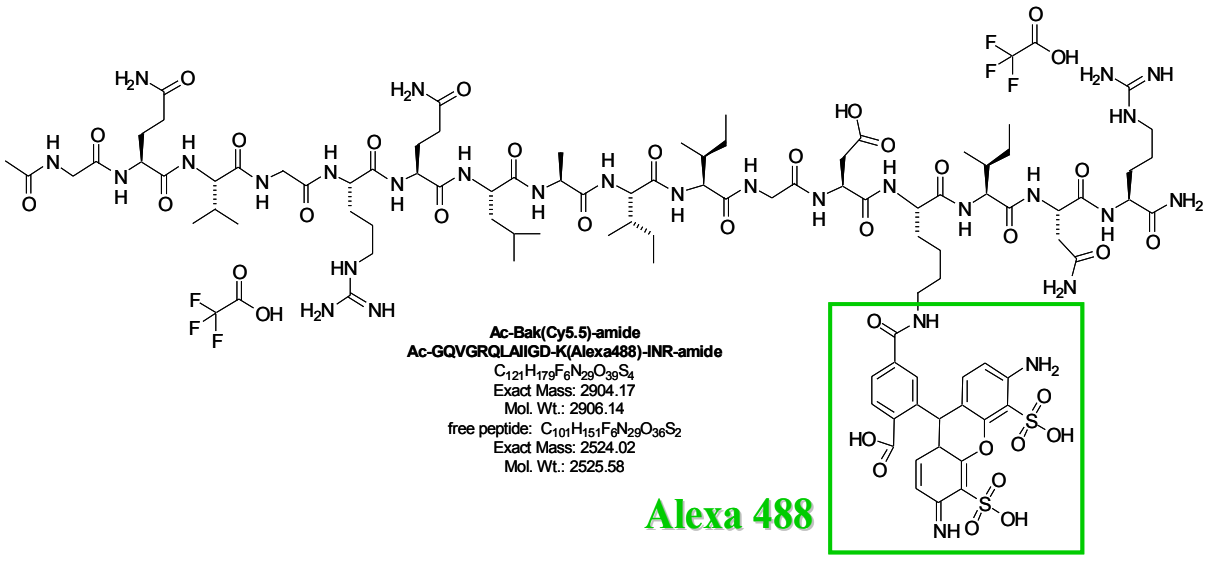


Figure 3: Fluorescently-labeled Bak peptide, which may be used in competitive binding assays with Bcl-xL

The fluorescently-labeled Bak peptide binds to the 20 kilo-Dalton Bcl-x_L protein and hence may be used in competitive binding assays of drug-like molecules that target Bcl-x_L. Figure 4 shows the results of a binding isotherm titration in which the total concentration of labeled Bak remains fixed at 1.1 nM. As the concentration of Bcl-x_L is increased, the labeled Bak becomes maximally bound and the steady state fluorescence anisotropy increases to an asymptotic value of 0.14. A curve fit to the data yields a dissociation constant of K_d = [Bcl-x_L][free ligand]/[complex] ≈ 2.3 nM, i.e., the ligand becomes 50% maximally bound for a Bcl-x_L concentration of about 2.3 nM.

Figure 5 shows an example of a competitive binding titration of an unspecified drug-like compound with Bcl-xL, again using FP as the method of assay. The total concentration of labeled Bak once more remains fixed at 1.1 nM, and the Bcl-x_L concentration is fixed at a nominal value of 2.4 nM, resulting in about 50% of maximal binding of the labeled Bak. Then, as the competitor drug-like compound concentration is increased, the complexed labeled-Bak becomes increasingly displaced, leading to a decrease in fluorescence anisotropy. A curve fit to Figure 5 yields a predicted Bcl-x_L concentration of 2.0 nM and a dissociation constant of K_d ≈ 0.18 nM.

The FP assays shown in Figures 4 and 5 were performed using 2 milliliter sample volumes with a measurement time of 15 seconds per data point using an Abbott clinical diagnostics instrument [9]. By contrast, FCS assays employ epi-illumination and confocal detection, and hence generally use much smaller microliter sample volumes. Figure 6 shows an example of a 0.1 s section of data with bursts of photons acquired using the Zeiss Confocor 2 FCS instrument, together with the normalized autocorrelation function (ACF) collected in a 30 second experiment. The graphs shown in Figure 6 were numerically calculated from the raw data file, which contains the times-of-arrival (TOAs) of each detected photon, as discussed in section 3 (see Figure 11). The graph of photon bursts exhibits both slow and fast bursts, corresponding to bound and unbound fluorescent ligands.

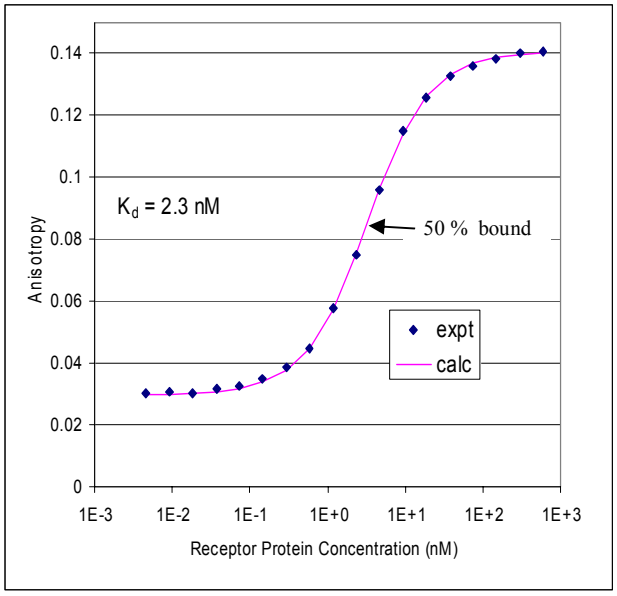


Figure 4: Binding isotherm titration of 1.1 nM Alexa-488-labeled Bak Vs.Bcl-x_L receptor protein, using FP assay.

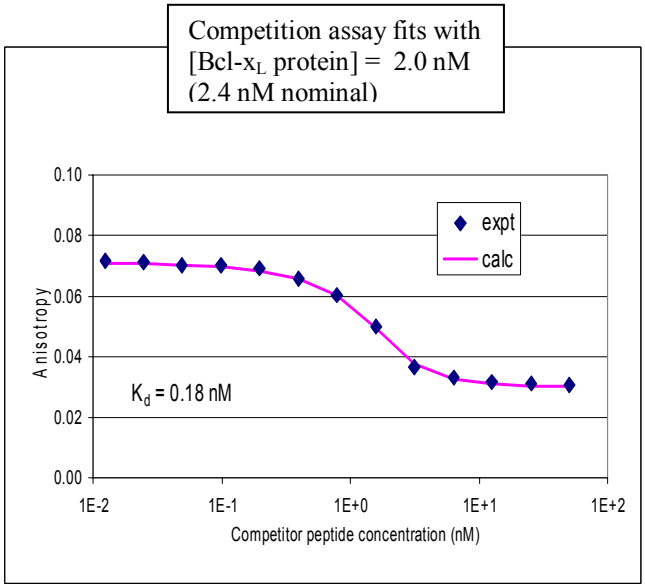


Figure 5: Example of competition assay titration of 1.1 nM Alexa-488-labeled Bak, nominally 2.4 nM of Bcl-x_L receptor protein, and increasing competitor.

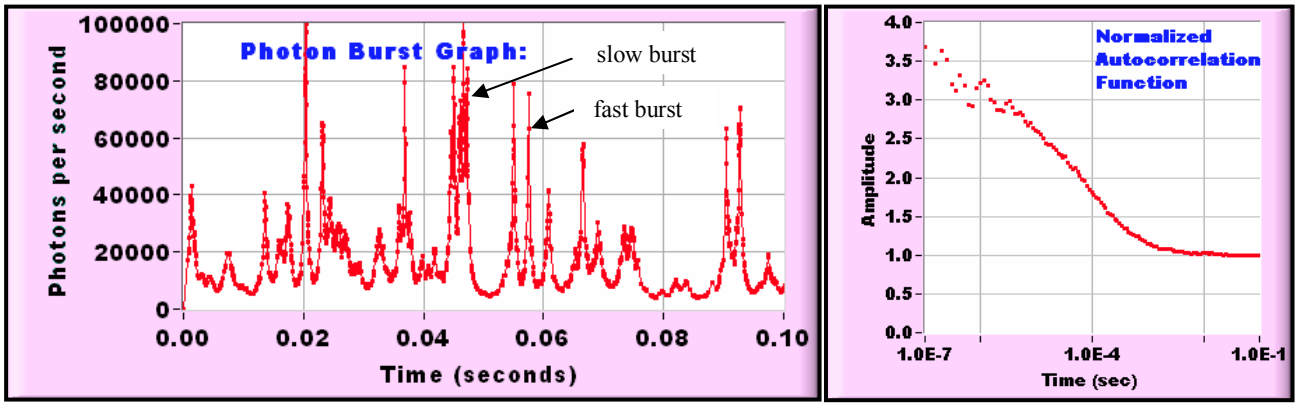


Figure 6: Section of photon burst data, and normalized autocorrelation function, from a sample at the 50% bound point in the titration curve of Figure 4, illustrating the use of FCS to assay ligand binding.

The normalized autocorrelation may be fit to a two-diffusion-component model to extract the relative fractions of bound and unbound ligands. For example, if the sample volume is taken to be a 3-D Gaussian cylindrical ellipsoid with semi-axes ω_0 and z_0, and if triplet kinetics are ignored, the normalized autocorrelation is of the classic form $g(\tau) = 1 + \gamma[\mathrm{L}(\tau/\tau_{D_1})/N_1 + \mathrm{L}(\tau/\tau_{D_2})/N_2]$, where $\mathrm{L}(\tau/\tau_i) = [1+\tau/\tau_i]^{-1}[1+\tau/\tau_i']^{-1/2}$ is a Lorentzian-shaped peak with characteristic width determined by $\tau_i = \omega_0^2/4D_i$ and $\tau_i' = z_0^2/4D_i$, D_i are the translational diffusion coefficients of bound and free fluorescently-labeled ligand, N_1 and N_2 are the mean numbers of bound and free label molecules within the sample volume (renormalized by the relative fluorescence brightness from bound and free labels), and γ is a constant dependent on the shape of the sample volume.

In this work, the normalized ACF was acquired using the Zeiss Confocor 2 with a 30 second acquisition time. The ACF was fit using the Zeiss Confocor 2 software to a two-diffusion component model, with fixed confocal parameter ω_0/z_0, fixed bound and free diffusional residence times, and an adjustable triplet component (which adds an exponential component to $g(\tau)$), as illustrated in the screen shot in Figure 7.

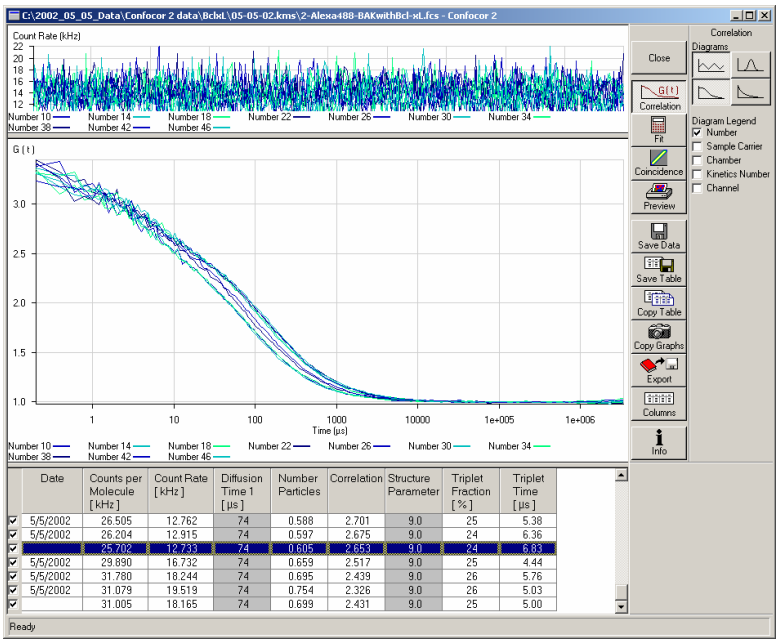


Figure 7: Screen shot of Zeiss Confocor 2 software for curve-fitting of normalized ACF.

Three independent FCS measurements, each for 30 seconds of acquisition time, were performed at each point of the binding isotherm titration of Figure 4. The percentage $N_1/(N_1+ N_2)\times100\%$ of the component with slower 146 µs diffusional residence time, corresponding to labeled Bak peptide bound to Bcl-x_L, as obtained from an unweighted nonlinear curve fit of the normalized ACF, is plotted in Figure 8. The total concentration of labeled Bak remains fixed at 1.1 nM and as the concentration of Bcl-x_L is increased, the labeled Bak becomes 100% bound. Also shown in Figure 8 is the total fluorescence count rate, which decreases by 12% as the labeled Bak becomes bound, indicating slight quenching of fluorescence in the bound configuration. Figure 9 shows data points obtained by averaging the 3 independent measurements, together with a curve fit to the binding isotherm, which yields a dissociation constant of $K_d \approx 2.3$ nM, in agreement with the result obtained by FP in Figure 4. (Note that the correction to K_d for the 12% quenching upon binding is small, as $N_1/(N_1+ 0.88\, N_2) = 0.53 \approx 0.5$ for $N_1 = N_2$.)

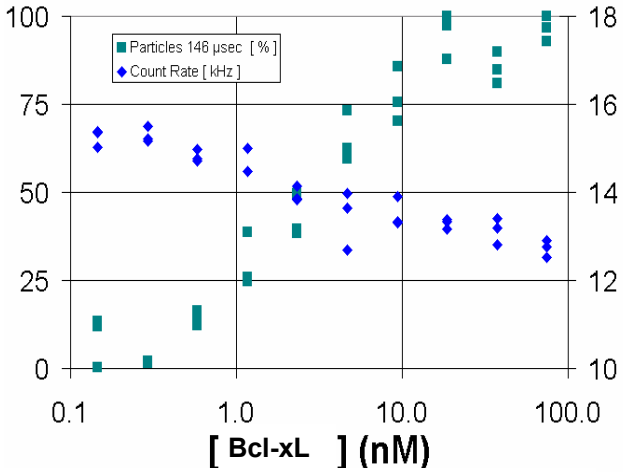


Figure 8: Binding isotherm titration of 1.1 nM Alexa-488-labeled Bak Vs. Bcl-x_L receptor protein, using FCS assay. The total count rate (right scale) decreases 12% due to quenching upon binding.

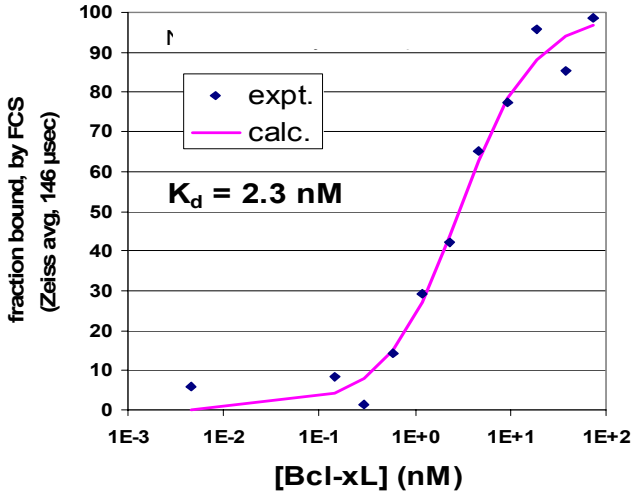


Figure 9: Curve fitting of the binding isotherm obtained by averaging the data in Figure 8 yields a dissociation constant of $K_d = 2.3$ nM.

3. MONTE CARLO SIMULATION

As discussed in section 1, systematic errors in the binding assay may arise from precipitation or aggregation of solution components, but curve fitting of the ACF in principle recovers the concentrations of labeled solution components through the dependence of the amplitudes of the ACF components on the mean numbers of labeled components within the sample volume, N_1 and N_2. However, background from autofluorescence of other solution components, instrument effects such as detector dead time, and photophysical effects such as excitation saturation and photobleaching may alter the ACF, leading to other systematic errors. Also, for high throughput screening, short duration data acquisition times will lead to random errors and possibly bias in the collected ACF. In order to evaluate short duration assay capabilities, to study the effects of realistic experimental conditions, and to facilitate comparison of different data analysis strategies, we have developed a sophisticated Monte Carlo simulation of single-molecule detection (SMD), FCS and fluorescence cross correlation spectroscopy (FCCS). Figure 10 shows a screen shot of the simulation during execution. The various simulation parameters for the optics, chromophore properties, and particle properties may be read in from a parameter file, and/or adjusted using the LabView controls on the three tab-panels at top left of the screen.

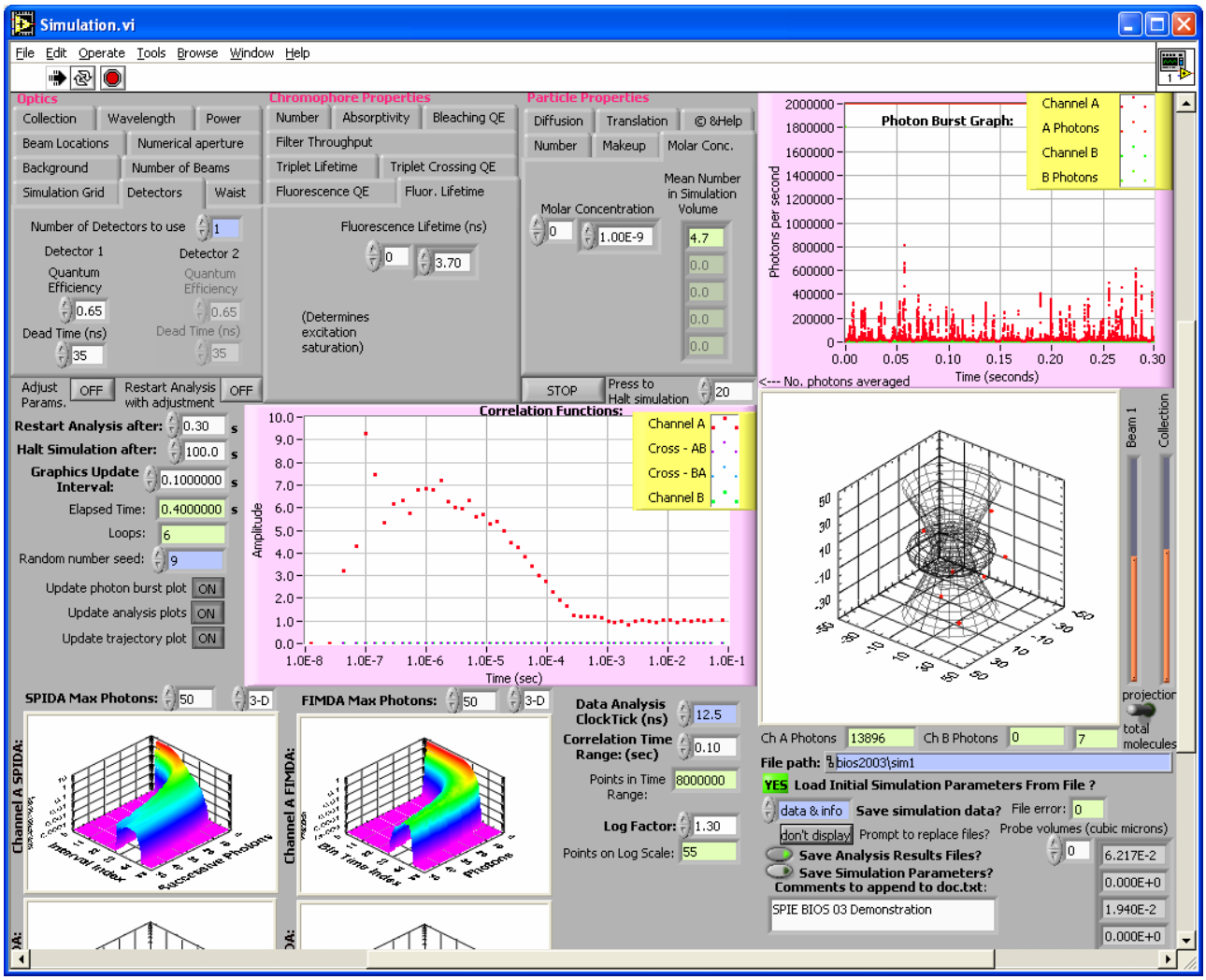


Figure 10: Screen shot of LabView interface to Monte Carlo simulation of FCS. The simulation volume, with focused laser beam, region of efficient light collection, and individual molecules, is shown in the lower right graph. The count rate averaged over 20 photons, with discrete photon bursts due to single molecules, is in the upper right graph. The normalized autocorrelation after an elapsed time of 0.4 s is in the center graph.

A key difference compared to prior simulations is that the new code generates the exact time of arrival (TOA) of each detected photon. Our previous simulation [10], which determined the photon TOAs to within 50 ns by use of a 50 ns time step, would take approximately 10 hours to simulate a 10 second experiment, and hence was executed on a 100-CPU Condor cluster. The new simulation uses a completely rewritten algorithm in which time increases continuously for all processes, but molecule locations are confined to a discrete grid. Molecules may enter or leave the simulation volume by diffusion or flow. The simulation can model one or two independently aligned laser excitation beams, one or two single photon detectors, several different chromophore types, and several particle types with different mobilities, labeled by specified numbers of the defined chromophore types. The execution time depends on input parameters and photon detection rate, but is of the order of 100 s for a 10 s experiment, including the data analysis described below. The code uses multiple threads, with key parts written in C and compiled using the Intel compiler for the Pentium IV.

4. DATA REDUCTION METHODS

Recording of the TOA of each detected photon provides the most complete experimental data acquisition modality for SMD and FCS. It was implemented in early SMD experiments using a 1 MHz analog-to-digital converter [11], and has been implemented in this work by access to the raw data files from the Zeiss Confocor 2 (which provides 20 MHz timing resolution), or by custom software with a National Instruments PCI-6602 counter/timer (80 MHz), or with a Picoquant Time Harp 200 (20 MHz). The various data reduction strategies may be applied by direct use of the TOAs with algorithms that provide rapid means of calculation. Figure 11 illustrates a simple algorithm by which the ACF may be accumulated in real time as a histogram of cumulative delay times [11], obviating the need for a hardware correlator. Reference [12] includes an algorithm for accumulation of cross-correlation functions (CCFs). Figure 11 also illustrates the accumulation of a family of histograms of delay times between successive numbers of photons (SPIDA) [10].

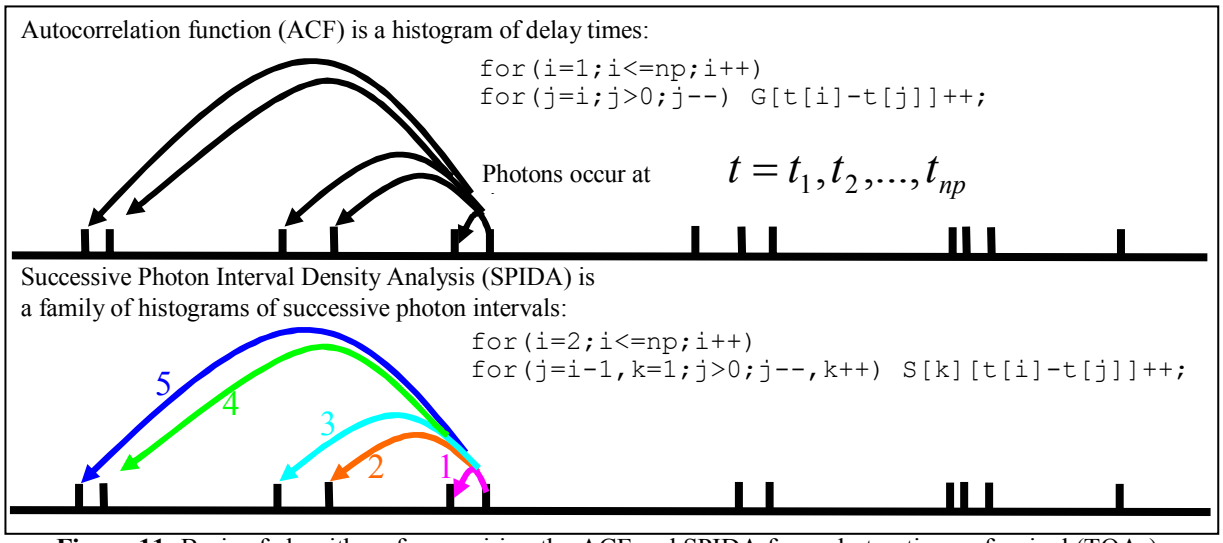


Figure 11: Basis of algorithms for acquiring the ACF and SPIDA from photon times of arrival (TOAs).

The normalized ACF contains information on the time scales of fluorescence fluctuations but does not capture information on the intensities of the fluctuations. Conversely, fluorescence intensity distribution analysis (FIDA), i.e., the histogram of photon counts within a bin of fixed width, captures information on intensities but not on the time scales. Figure 12a illustrates the FIDA from simulations in which histograms are obtained using a 100 μs sliding bin. In this example, the sample assay would be achieved by resolving the histogram labeled "BG+0.1 nM of each" into the components due to background, fluorophore with quantum efficiency QE=0.2, and fluorophore with QE=0.9. Figure 12b illustrates fluorescence intensity multiple density analysis (FIMDA) [13], i.e., the FIDA obtained with a complete range of bin widths, with selectable logarithmically-spaced widths, as illustrated in Figure 12c. FIMDA and SPIDA capture information on both the time scales and intensities of fluorescence fluctuations. We have developed a fast algorithm for obtaining FIMDA directly from the TOAs, so that the accumulation of the FIMDA and SPIDA, (as well as ACF, and CCF) are all obtained during the simulation in real time, as illustrated in the lower left graphs in Figure 10.

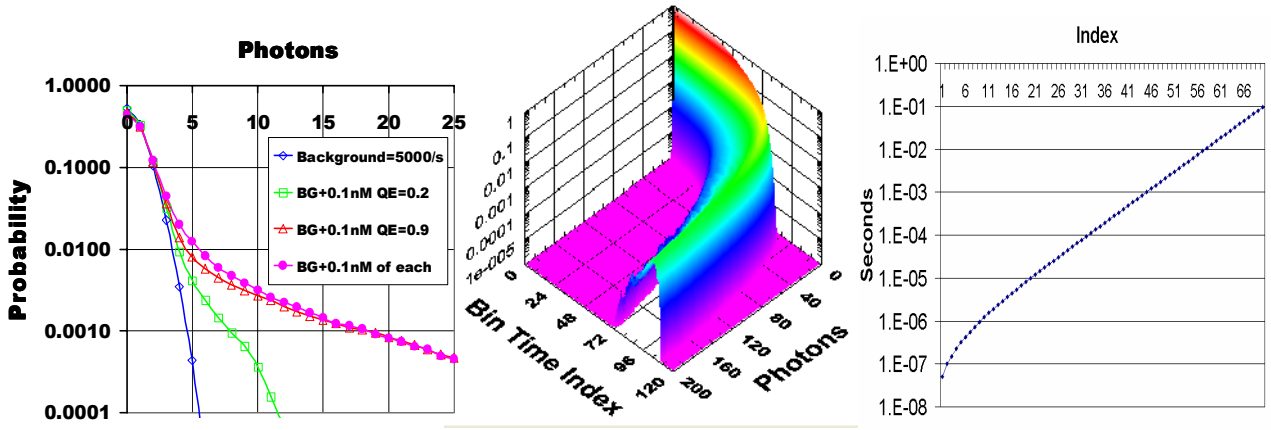


Figure 12: (a) FIDA for a 100 μs bin width; **(b)** FIMDA or family of FIDA histograms for selected bin widths; **(c)** Bin width in seconds Vs. Bin Time Index of FIMDA-curves illustrating the logarithmic spacing of selected bin widths.

5. EXAMPLES OF SIMULATION RESULTS

To evaluate how quantitative the simulation predictions are, the simulation was executed for a contrived set of ideal parameters to model a nanomolar concentration of an ideal fluorophore and a spherical 3-D Gaussian excitation and collection volume, with zero triplet crossing, photobleaching, excitation saturation, detector dead time, and background. Curve fitting of the resultant normalized ACF from a 100 s simulation to the one component classic form introduced in section 2, $g(\tau) = 1 + \gamma \mathrm{L}(\tau / \tau_{D_1}) / N_1$, gives values for the amplitude and width that agree within about 5% with theoretical predictions, shown by the solid line in Figure 13a, obtained by using numerical integration to evaluate $\gamma = \int I^2(r) C^2(r) dr^3 \Big/ \int I(r) C(r) dr^3$, where *I(r)* and *C(r)* represent the excitation intensity and collection efficiencies normalized to their peak values at the origin [14]. However, if the extent of the simulation grid is reduced below about $5\omega_0$, the tail of the ACF peak is altered, because molecules that leave and then reenter the laser focal volume to give long-time correlations are not correctly modeled. Also, as the simulation grid resolution is reduced below about $\omega_0/5$, the mean number of molecules in the laser focal volume N_1 becomes inaccurate and the amplitude of the normalized ACF becomes smaller than expected.

The simulation was used to study the effect on the ACF of background counts, as may arise from autofluorescence from sample constituents. Figure 13b shows that a modest background count rate will lead to a decreased amplitude and hence an overestimate of the fluorophore concentration. Figure 13c shows the effects of increasing laser power on the ACF due

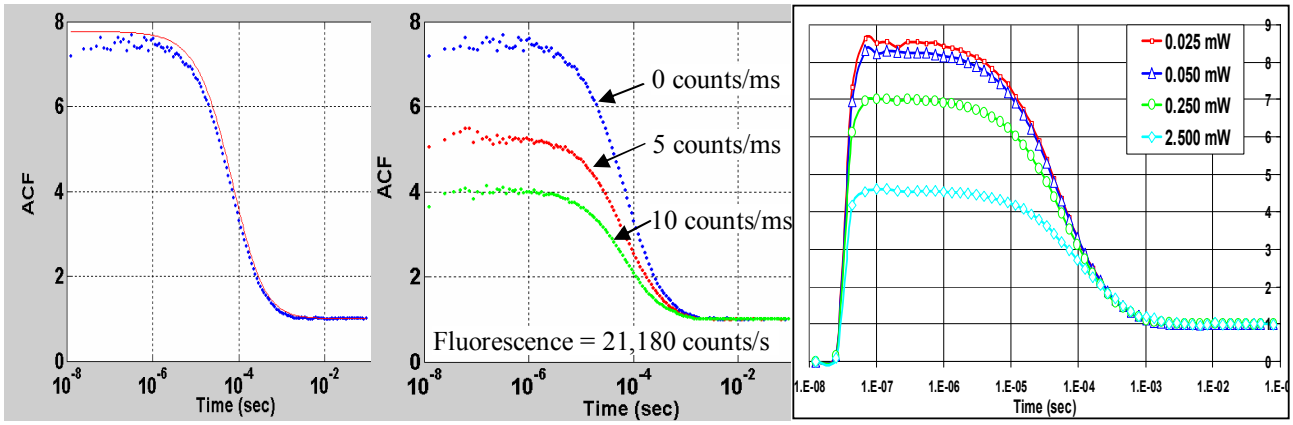


Figure 13: (a) Normalized ACF and theoretical prediction (solid line) for contrived ideal conditions, discussed in text; **(b)** Effect of background on the amplitude of the normalized ACF; **(c)** Effect of laser power on the normalized ACF due to saturation, with Zeiss Confocor 2 geometry and conditions of Figure 10.

to excitation saturation, in the case of the Zeiss Confocor 2 geometry [15]. The truncation of the ACFs at short delays due to the 35 ns detector dead time is also evident in Figure 13c. In the simulation, saturation is modeled by causing excited molecules to wait a random time with mean equal to their specified fluorescence lifetime before they may be excited again, thereby limiting their overall rate of excitation. Berland has reported a heuristic model of saturation effects in the case of FCS with 2-photon excitation [16] and it is no surprise that similar distortion of the ACF is seen here in the 1-photon case, although the relatively low power level at which saturation begins to give noticeable effects may be somewhat surprising. It is clear that the distortion in the ACF is due to saturation as the simulation allows other effects, such as increased background with laser power, and triplet crossing, to be switched off.

The simulation is useful for investigating a host of experimental configurations and comparing different data analysis strategies. Here we compare different methods for analyzing the data in a 2-color 2-detection channel experiment. In such experiments, the cross correlation function (CCF) is usually used to detect the presence of particles simultaneously labeled by both fluorophores. However, the CCF must be corrected for spectral cross talk, and the corrected amplitude is usually small and noisy, dictating prolonged data acquisition times. Would it be better to search through the two data streams for simultaneously occurring photon bursts? As diffusing particles may give rise to photon bursts of different durations and amplitudes, we chose to use the entire range of possible bin widths, as shown in the FIMDA analysis of Figure 12b. The data analysis is fast when using algorithms that directly use the TOAs. Indeed, we find an assay based on counting simultaneous bursts is superior to one based on the CCF. Figures 14a and b show two cases of 2-dimensional histograms of the numbers of photons within sliding bins of 35 μs width, using data generated by 1s simulations. In the second case, particles labeled by both fluorophores, which give rise to simultaneous photon bursts of similar amplitudes, are clearly present, but the change in the amplitude of the CCF is small and is compounded by cross-talk, as seen by comparing Figures 14c and d. Correction for cross talk requires measure of the individual concentrations from the ACFs and knowledge of the extent of the cross-talk [17]. Also, note that the simultaneous bursts are not quite so clearly evident when using longer or shorter bin widths, or when collecting 2-D histograms using jumping rather than sliding bins, as is the usual approach in 2-D FIDA [18].

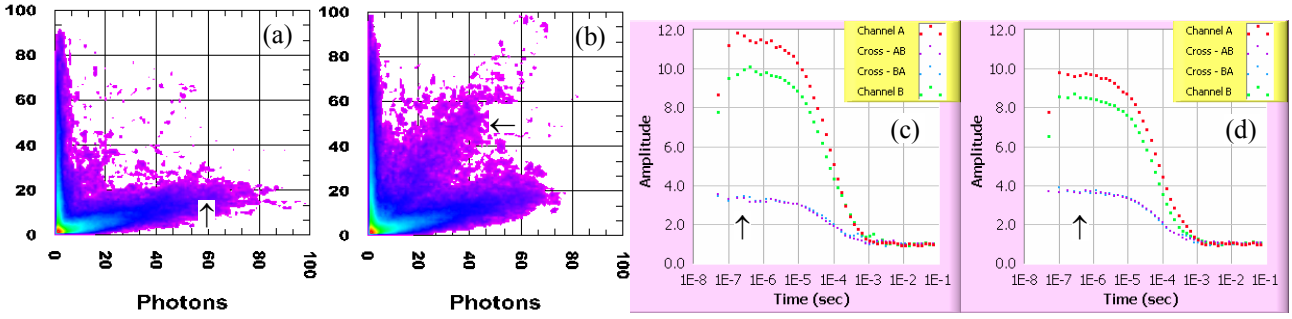


Figure 14: (a) One from a family of graphs of 2-D FIMDA, here with sliding bin of 35 μs width, from a mix of two differently colored dyes each at 4×10^{-10}M, with 20% spectral cross-talk, causing photons from one dye to appear partly in the other detector channel (↑). **(b)** The same as (a) but from a mix of 5×10^{-10}M of each dye, and also with 5×10^{-11}M of particles labeled by both dyes, which yield clearly distinguishable simultaneous bursts of similar amplitudes (←). **(c)** The correlation functions from the data of (a), with the CCFs indicated (↑). **(d)** The correlation functions from the data of (b), with little change in the CCFs (↑) compared to (c).

6. ERRORS IN THE ACF FOR SHORT DURATION COLLECTION TIMES

Whereas the definition of the ACF involves an integral over an infinite duration, it is to be understood that evaluations based on finite duration experiments are only estimates of the true ACF, just as a sample mean is merely an estimate of the population mean. From the earliest days of FCS, the estimate of the normalized ACF from finite data sets was reported to be biased [19, 20] and much effort was made to determine the statistical error bars at points in the normalized ACF. Here we compare the errors and bias using several different ways of estimating the normalized ACF from short duration finite data sets, which have been generated by the simulation to avoid experimental effects such as laser drift.

Figure 15 illustrates 4 formulae for evaluating the normalized ACF, which we have implemented directly from the photon TOAs. In $g_0(\tau)$, there is an attempt to correlate data with other data outside the range of the experiment, similar to the "mistakes" in the equations (4) of Refs. [19] and [20]. Hardware correlators may possibly be prone to such a mistake. The evaluation still works if the out-of-range data is assumed to take some value, typically zero. In $g_1(\tau)$, it is recognized that the ACF is to be evaluated out to a time τ = T, and hence accumulation of the ACF waits until after a duration T of the data has passed. This is the method usually used in our simulation to accumulate the ACF (e.g., in Figure 10, center), because of its simplicity. However, it does not efficiently utilize the data. In $g_2(\tau)$, different wait times τ are used before evaluation of each point in the ACF. Also, different values of the summed intensity must be used to normalize the ACF at each point τ. In $g_3(\tau)$, symmetrical normalization is used, thereby requiring three different summations for each point τ.

$$g_0(\tau) = \frac{T\sum_{t=0}^{T} I(t)I(t-\tau)}{\left[\sum_{t=0}^{T} I(t)\right]^2}, \tau = 0,..., \mathrm{T}$$

$$g_1(\tau) = \frac{(T-\mathrm{T})\sum_{t=\mathrm{T}}^{T} I(t)I(t-\tau)}{\left[\sum_{t=\mathrm{T}}^{T} I(t)\right]^2}, \tau = 0,..., \mathrm{T}$$

$$g_2(\tau) = \frac{(T-\tau)\sum_{t=\tau}^{T} I(t)I(t-\tau)}{\left[\sum_{t=\tau}^{T} I(t)\right]^2}, \tau = 0,..., \mathrm{T}$$

$$g_3(\tau) = \frac{(T-\tau)\sum_{t=\tau}^{T} I(t)I(t-\tau)}{\left[\sum_{t=\tau}^{T} I(t)\right]\left[\sum_{t=\tau}^{T} I(t-\tau)\right]}, \tau = 0,..., \mathrm{T}$$

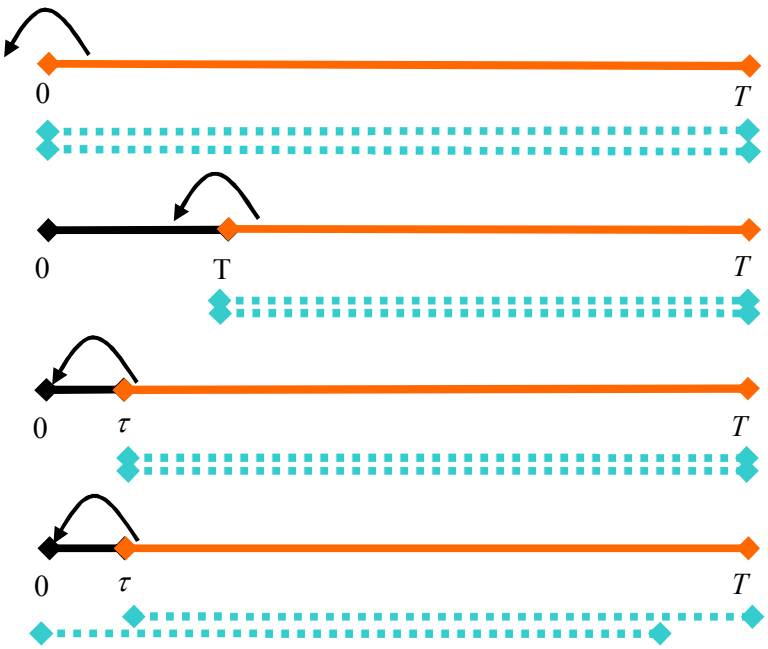


Figure 15: Comparison of four methods for estimation of the ACF from finite duration data sets. The solid lines at right illustrate the time range of the summations in the numerators of the equations at left. The grey dashed lines illustrate the time ranges of the summations in the denominators, used to normalize the ACF.

For 0.2 s simulations of sulforhodamine 101 in water with the Confocor 2 geometry of Figure 10 and evaluation of the ACF out to τ = 0.1 s, we find that g_0 yields biased estimates of the ACF in that the average of 100 ACFs from different 0.2 s simulations does not approach the ACF from a 100 s simulation, particularly at the longer dwell times, as seen in Figure 16a. We find g_1 yields unbiased estimates, but the spread in ACFs is considerably greater, as seen in Figure 16b. In particular, the arrows in Figure 16b indicate considerable uncertainty in the width of the ACF. Note that the errors at nearby points in each ACF estimate are correlated, so that the ACF from an individual 0.2 s simulation may have faster or slower width, but on average is unbiased. The fact that errors at nearby points in the ACF are not independent but are correlated means that efforts to determine the error bars at each point in the ACF for use in weighted least-squares fitting of the ACF will never provide a statistically rigorous means of analysis, because least-squares curve fitting implicitly requires each data point to be statistically independent. We find g_2 and g_3 also yield unbiased estimates, but they exhibit considerably less fluctuations in their widths than g_1, as may be expected from the more efficient use of data. We find that the symmetrical normalization in g_3 yields smaller errors than g_2 in the tail of the ACF, at the arrow of Figure 16c, as has been reported elsewhere [21]. Also, by comparing Figures 16c and 16d, we see that evaluating the ACF at fewer points by averaging the g_3 values at a larger numbers of adjacent τ values, corresponding to a larger log factor and steeper slope in Figure 12c, leads to less point-to-point fluctuations in an individual ACF. Finally, Figure 17 shows that contrary to the case of the multi-tau correlator discussed in Reference [22], the g_3 estimate of the ACF from 0.2 s duration experiments remains unbiased, being within the random errors exhibited by the ACFs from several different 20 s simulations, even for a slow diffusion coefficient of 4×10^{-7} cm^2/s.

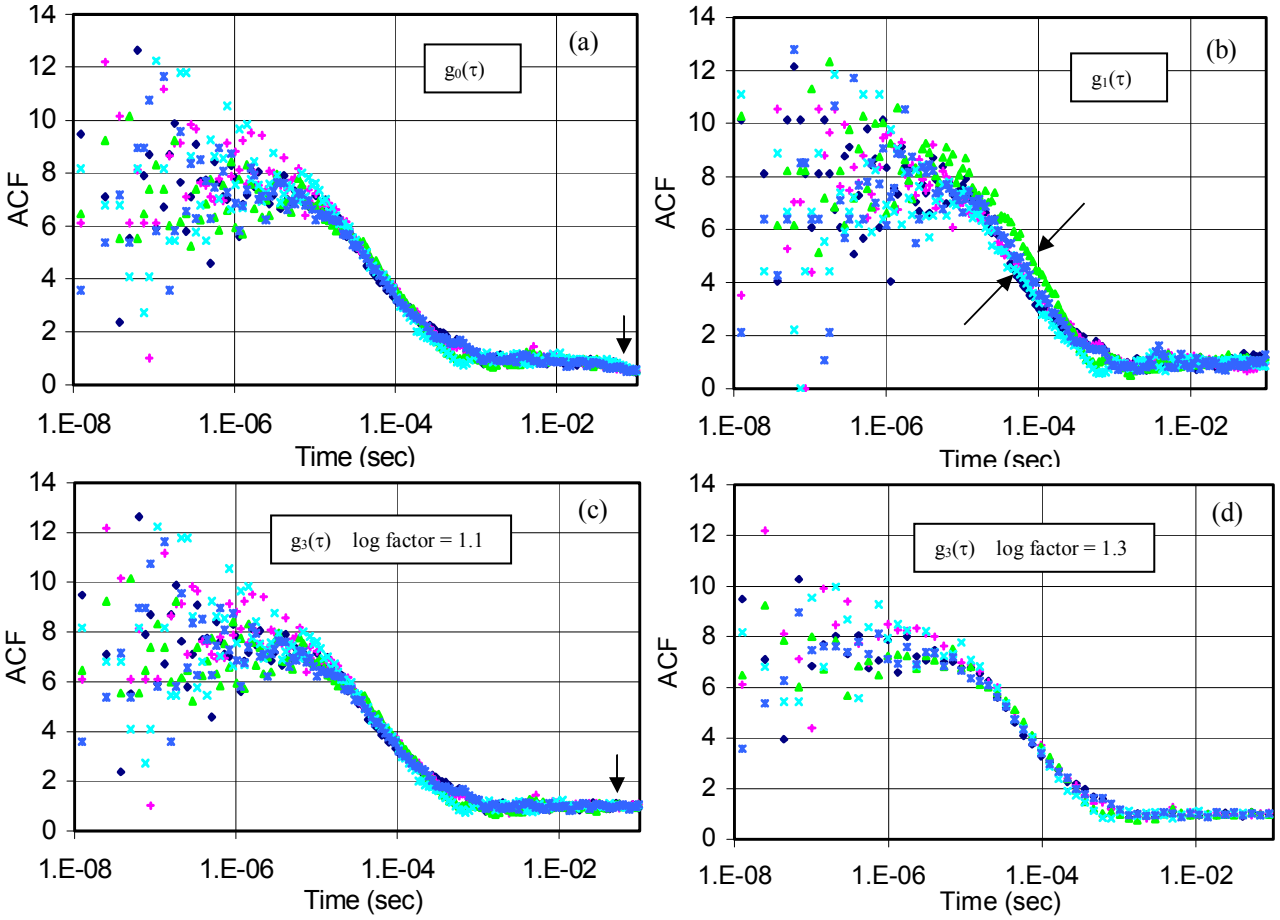
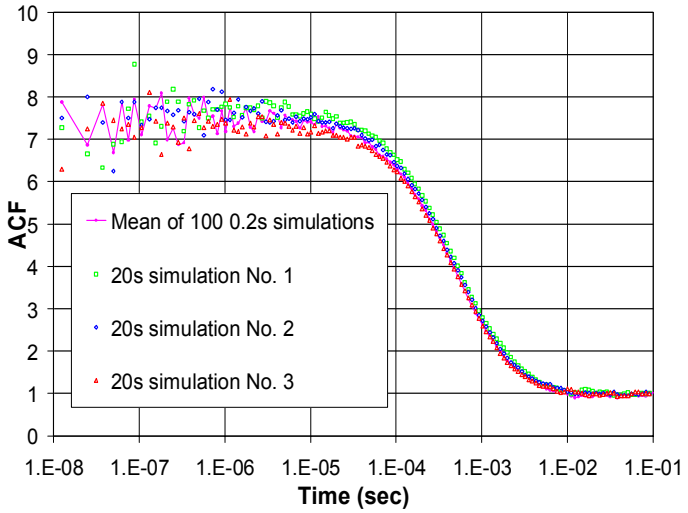


Figure 16: Comparison of the ACF obtained from the same set of five different 0.2s simulations, as evaluated by different methods discussed in the text.



Simulation conditions: parameters similar to those of Reference [22] (but with only a 0.2 s rather than 3-30 s experiment duration), namely a slow diffusion coefficient of $D = 4 \times 10^{-7}$ cm^2/s (corresponding to 2.5 MDa dextran molecules, in our case labeled by sulforhodamine 101), a laser beam waist of 0.5 μm (yielding a calculated diffusional residence time of $\tau = \omega_0^2/4D = 3125$ μs), low laser power (20 μW) to avoid saturation, resulting in a fluorescence count rate of about 2800 counts/s, no triplet crossing, bleaching, background, or detector dead time. The normalized ACF estimate obtained from such a finite duration experiment with a multi-tau correlator would exhibit bias according to the "phase diagram" of Figure 5 of Reference [22].

Figure 17: The g_3 estimates of the ACF from 0.2 s simulations are unbiased, as seen by comparing the mean of 100 such simulations with several different 20 s simulations. The lack of bias holds even for a slow diffusion coefficient, which results in marked variation of the photon bursts, photon counts and ACF generated by different 0.2 s simulations.

7. REFERENCES

1. K. Swift, S.N. Anderson, and E. Matayoshi, "Dual laser fluorescence correlation spectroscopy as a biophysical probe of binding interactions: Evaluation of new red-fluorescent dyes," *Proc. of SPIE: Advances in Fluorescence Sensing Technology V,* J.R. Lakowicz, R.B. Thompson, Eds., **4252**, pp. 47-58 (2001).
2. K. Swift, and E. Matayoshi, "Evaluation of a red fluorescent dye for high throughput compound screening applications in pharmaceutical discovery research," *Proc. of SPIE: Advances in Fluorescence Sensing Technology IV,* J.R. Lakowicz, S.A. Soper, R.B. Thompson, Eds., **3602**, pp. 68-74 (1999).
3. K. Swift, and E. Matayoshi, "High throughput screening using dynamic fluorescence," *Proc. of SPIE: Advances in Fluorescence Sensing Technology II,* J.R. Lakowicz, Ed., **2388**, pp. 182-189 (1995).
4. E. Matayoshi and K. Swift, "Applications of FCS to protein-ligand interactions: Comparison with Fluorescence Polarization," pp. 84-98 in *Fluorescence Correlation Spectroscopy: Theory and Applications*, R. Rigler, E.S. Elson, Eds., Springer-Verlag Berlin Heidelberg, IBBN 3-540-67433-0 (2001).
5. Y. Chan, J.D. Muller, P.T.C. So, and E. Gratton, "The photon counting histogram in fluorescence fluctuation spectroscopy," *Biophys. J.*, **77**, pp. 553-567 (1999).
6. P. Kask, K. Palo, D. Ullmann, and K. Gall, "Fluorescence-intensity distribution analysis and its applications in biomolecular detection and technology," *Proc. Natl. Acad. Sci. USA* **6**, pp. 13756-13761 (1999).
7. S. W. Fesik and Y. Shi, "Controlling the caspases," *Science*, **294**, pp. 1477-1478 (2001).
8. M. Sattler, H. Liang, D. Nettesheim, R. P. Meadows, J. E. Harlan, M. Eberstadt, H. S. Yoon, S. B. Shuker, B. S. Chang, A. J. Minn, C. B. Thompson, S. W. Fesik, "Structure of Bcl-x_L-Bax peptide complex: recognition between regulators of apoptosis," *Science*, **275**, pp. 983-986 (1997).
9. A.M. Petros, D.G. Nettesheim, Y. Wang, E.T. Olejniczak, R.P. Meadows, J. Mack, K. Swift, E.D. Matayoshi, H. Zhang, C.B. Thompson, and S.W. Fesik, "Rational for Bcl-x_L/Bad peptide complex formation from structure, mutagenesis, and biophysical studies," *Protein Science*, **9**, pp. 2528-2534 (2000).
10. L.M. Davis, P.E. Williams, H.M. Cain, D.A. Ball, C.G. Parigger, E.D. Matayoshi, and K.M. Swift, "Comparison of fluorescence correlation spectroscopy and other single-molecule data analysis methods for assay of protein-ligand interactions," 46-th Annual Meeting of the Biophysical Society, San Francisco, California, Feb. 23-27, 2002; *Biophys. J.*, **82**, p. 43a (2002).
11. L.Q. Li and L.M. Davis, "Rapid and efficient detection of single chromophore molecules in aqueous solution" *Applied Optics*, **34**, 3208 (1995).
12. L.M. Davis, J.G.K. Williams, and D.T. Lamb, "Computer simulation of gene detection without PCR by single molecule detection," *Proc. of SPIE: Biomedical Sensors, Fibers, and Optical Delivery Systems*, F. Baldini, N.I. Croitoru, M. Frenz, I. Lundstrom, M. Miyagi, R. Pratesi, O.S. Wolfbeis, Eds., **3570**, pp. 282-293 (1999).
13. K. Palo, U. Mets, S. Jager, P. Kask, and K. Gall, "Fluorescence intensity multiple distributions analysis: Concurrent determination of diffusion times and molecular brightness," *Biophys. J.*, **79**, pp. 2858-2866 (2000).
14. T. Wohland, R. Rigler, H. Vogel, "The standard deviation in fluorescence correlation spectroscopy," *Biophys. J.*, **80**, pp. 2987-2999 (2001).
15. T. Jankowski, and R. Janka, "Confocor 2: The second generation of fluorescence correlation microscopes," pp. 331-345 in *ibidem* [4].
16. K. Berland, and G. Shen, "Investigating two-photon photophysics with fluorescence correlation spectroscopy," *Proc. of SPIE: Multiphoton Microscopy in the Biomedical Sciences III*, A. Perisamy, P.T.C. So, Eds., **4963**, paper 04 (2003).
17. P. Schwille, "Cross-correlation analysis in FCS", pp. 360-378 in *ibidem* [4].
18. P. Kask, K. Palo, N. Fay, L. Brand, U. Mets, D. Ullmann, J. Jungmann, J. Pschorr, and K. Gall, "Two-dimensional Fluorescence Intensity Distribution Analysis: Theory and Applications," *Biophys. J.*, **78**, pp. 1703-1713 (2000).
19. E. Jakeman, E.R. Pike, and S. Swain, "Statistical accuracy in the digital autocorrelation of photon counting fluctuations," *J. Phys. A: Gen. Phys.*, **4**, pp. 517-534 (1971).
20. D.E. Koppel, "Statistical accuracy in fluorescence correlation spectroscopy," *Phys. Rev. A*, **10**, pp. 1938-1945 (1974).
21. K. Schatzel, M. Drewel, and S. Stimac, "Photon correlation measurements at large lag times: improving statistical accuracy," *J. Mod. Optics*, **35**, pp. 711-718 (1988).
22. S. Saffarian and E. Elson, The statistical analysis of fluorescence correlation spectroscopy: standard deviation and bias," *Biophys. J.*, **84**, pp. 2030-2042 (2003).