Focal Volume Optics and Experimental Artifacts in Confocal Fluorescence Correlation Spectroscopy

Samuel T. Hess* and Watt W. Webb[†]

*Department of Physics and [†]School of Applied and Engineering Physics, Cornell University, Ithaca, New York 14853 USA

ABSTRACT   Fluorescence correlation spectroscopy (FCS) can provide a wealth of information about biological and chemical systems on a broad range of time scales (<1 μs to >1 s). Numerical modeling of the FCS observation volume combined with measurements has revealed, however, that the standard assumption of a three-dimensional Gaussian FCS observation volume is not a valid approximation under many common measurement conditions. As a result, the FCS autocorrelation will contain significant, systematic artifacts that are most severe with confocal optics when using a large detector aperture and aperture-limited illumination. These optical artifacts manifest themselves in the fluorescence correlation as an apparent additional exponential component or diffusing species with significant ($>30\%$) amplitude that can imply extraneous kinetics, shift the measured diffusion time by as much as $\sim80\%$, and cause the axial ratio to diverge. Artifacts can be minimized or virtually eliminated by using a small confocal detector aperture, underfilled objective back-aperture, or two-photon excitation. However, using a detector aperture that is smaller or larger than the optimal value (~4.5 optical units) greatly reduces both the count rate per molecule and the signal-to-noise ratio. Thus, there is a tradeoff between optimizing signal-to-noise and reducing experimental artifacts in one-photon FCS.

INTRODUCTION

Fluorescence correlation spectroscopy (FCS) (Magde et al., 1972, 1974; Elson and Magde, 1974) is an elegant and sensitive technique for measuring dynamic processes that manifest themselves in a fluorescent signal on the submicrosecond-to-second time scales. With origins in quasielastic light scattering (Cummins and Swinney, 1970) and a basis in the statistical thermodynamics of fluctuations in solution, FCS has been used to measure diffusion coefficients, chemical kinetics, excited-state molecular dynamics, picomolar concentrations (Eigen and Rigler, 1994), and the dynamics of the interaction of fluorescent molecules in vitro (Koppel et al., 1976; Magde et al., 1972). Recently there has been rapidly increasing use of FCS for biological applications (Hess et al., 2002; Rigler and Elson, 2001; Schwille et al., 1996; Maiti et al., 1997) including living biological systems (Schwille et al., 1999a; Brock et al., 1998, 1999; Hink et al., 2000; Wachsmuth et al., 2000; Cluzel et al., 2000; Politz et al., 1998).

What are the best optical operating conditions for FCS? That question comprises the focus of this study. Many FCS systems allow changes in the type and numerical aperture (NA) of the objective lens, the size of the detector aperture, and the degree of underfilling of the back-aperture of the objective. Previous theoretical work has investigated the signal-to-noise ratio (S/N) for FCS systems (Koppel, 1974; Qian and Elson, 1991), and confocal microscopes (Gu and Gan, 1996; Sheppard et al., 1991; Sandison and Webb, 1994; Gu and Sheppard, 1993), but most previous work has made significant approximations in the treatment of the focal volume (Rigler et al., 1993), particularly at high-NA (Sheppard and Matthews, 1987) and in the treatment of underfilling, except in the case of two-photon excitation (W. Zipfel, personal communication). It is commonly assumed that the observation volume in FCS is a Gaussian in three dimensions. However, it will be shown that this is not sufficiently accurate under many measurement conditions and may lead to inaccurate results. To attempt to improve upon current FCS methodology we use a more sophisticated description of the focal volume (Wolf, 1959; Richards and Wolf, 1959), which treats the polarization of the excitation and is well-suited to high-NA optics. Furthermore, we present theoretical predictions and measurements together. We also attempt to determine the values of the detector aperture and underfilling fraction that yield the highest S/N and minimize artifacts in the FCS autocorrelation function that result from the non-Gaussian nature of the observation volume.

THEORY

Although FCS is an excellent technique for extracting a variety of quantitative information from biological systems, the influence of the illumination and collection optics on the measured autocorrelation must be considered to ensure accurate results. This section delineates the relationship between the spatial profile of the focal volume and parameters measured by FCS.

Introduction to FCS

Fluorescence correlation spectroscopy (Magde et al., 1972), which is based on fluctuation analysis of the fluorescence from an observation volume on the order of a femtoliter

containing a small number ($<$1000) of molecules, provides quantitative physical and chemical kinetic information on time scales from 10^{-7} to $>10^2$ s. FCS has been demonstrated to be quite useful for photophysical characterization of sparse fluorescent molecules (Mertz et al., 1995) and measurement of dynamics of those molecules that give rise to fluctuations in their fluorescence (Schwille et al., 1996; Heikel et al., 2000). Technological innovations, adapted particularly in the Rigler and Eigen groups (Eigen and Rigler, 1994), have resulted in a recent revival of the technique, the fundamentals of which were first developed in the early 1970s in the Webb group. The basis of the technique is the observation of the fluorescence F, produced by dilute fluorescent species ($\sim$nM concentrations) that diffuse and react chemically according to

$$\frac{\partial}{\partial t} \delta C_j(\mathbf{r}, t) = D_j \nabla^2 \delta C_j(\mathbf{r}, t) + \sum_k T_{jk} \delta C_k(\mathbf{r}, t) \quad (1)$$

where the concentration fluctuation δC_j of the jth species from the mean (temporal average) $\langle C_j \rangle$ is given by $\delta C_j(\mathbf{r}, t) = C_j(\mathbf{r}, t) - \langle C_j(\mathbf{r}, t) \rangle$, as a function of time (t), the diffusion coefficient D_j, and the tensor T_{jk}, which expresses the chemical reaction kinetics and stoichiometry between species.

The concentration fluctuations result in fluorescence fluctuations $\delta F(t)$, which are related to the experimental illumination and collection profiles (now considering a single species) by

$$\delta F(t) = I_O(\mathbf{r}_O = 0) \int S(\mathbf{r}_O) \Omega(\mathbf{r}_O) \delta(q\sigma C(\mathbf{r}_O, t)) d\mathbf{r}_O \quad (2)$$

where I_O is the illumination intensity profile, $S \equiv I(\mathbf{r}_O)/I(\mathbf{r}_O = 0)$ is the normalized illumination profile, Ω is the collection efficiency profile, $\mathbf{r}_O$ is position in object space, and q, σ, and C are the effective quantum yield, extinction coefficient, and concentration, respectively, of the observed fluorescent species. Clearly based on Eq. 2, the spatial dependence of the product of S and Ω, the normalized observation volume profile

$$O(\mathbf{r}_O) \equiv S(\mathbf{r}_O)\Omega(\mathbf{r}_O), \quad (3)$$

is crucial in determining the spatial distribution of fluorescence fluctuations that are to be observed. In practice, kinetic information is extracted from the fluorescence fluctuations by autocorrelation (Webb, 1976):

$$G(\tau) = \frac{\langle \delta F(t) \delta F(t + \tau) \rangle}{\langle F(t) \rangle^2} \quad (4)$$

where $\delta F(t) = F(t) - \langle F(t) \rangle$, and τ is time delay. $G(\tau)$ contains diffusion, chemical kinetic, molecular brightness, concentration, and photophysical information by quantify-

ing the temporal decay (as a function of τ) of the fluorescence fluctuations.

A solution to the autocorrelation function may be formulated in the case of ideal solutions and concentration- and position-independent diffusion coefficients using Green's function for classical diffusion kinetics:

$$\Psi(\mathbf{r}) = \frac{N}{(4\pi D\tau)^{3/2}} \exp\left[-\frac{|\mathbf{r}|^2}{4D\tau}\right] \quad (5)$$

where τ is time, N is the number of particles, D is the diffusion coefficient, and $\mathbf{r}$ is position. The diffusion autocorrelation function can be calculated using a double integral:

$$G_D(\tau) = \int\int d\mathbf{r}_1 d\mathbf{r}_2 O(\mathbf{r}_1) \Psi(\mathbf{r}_1 - \mathbf{r}_2, \tau) O(\mathbf{r}_2). \quad (6)$$

Autocorrelation function assuming a Gaussian observation volume

The standard autocorrelation function for a single fluorescent species freely diffusing in an ellipsoidal 3D-Gaussian observation volume $O(\mathbf{r}_O)$ at low intensity (well below saturation) has the solution (Schwille et al., 1996):

$$G_D(\tau) = \frac{1}{N} \cdot \left[1 + \frac{\tau}{\tau_D}\right]^{-1} \left[1 + \frac{\tau}{\omega^2 \tau_D}\right]^{-0.5} \quad (7)$$

where τ_D is the characteristic (diffusion) time molecules spend on the average in the observation volume, ω is the axial ratio (ratio of axial to radial dimensions of the observation volume), and N is the average number of molecules. Autocorrelation and volume are clearly related in the limit $\tau \to 0$, where

$$G_D(\tau \to 0) = 1/N = 1/C_0 V \quad (8)$$

and $C(\mathbf{r}) = C_0$ is a spatially constant concentration of molecules in the observation volume of size V. Note that both x and y (z is axial) contribute a factor of $(1 + D\tau/\rho^2)^{-0.5}$ to $G_D(\tau)$, where ρ is the transverse $1/e^2$ radius of $O(\mathbf{r}_O)$, but that Eq. 7 is the special case where $\rho_x = \rho_y$, and hence only two diffusion time constants exist: $\tau_D \equiv \rho^2/4D$, $\tau' = (\omega\rho)^2/4D$.

For anomalous diffusion, where $\langle r^2 \rangle = \Gamma t^\alpha$ and $4Dt \to \Gamma t^\alpha$, the standard fitting function is (Schwille et al., 1999b)

$$G(\tau) = \frac{1}{N} \cdot \frac{1}{1 + \Gamma \tau^\alpha/\rho^2} \quad (9)$$

which also assumes a Gaussian observation volume, where the temporal exponent α has the possible range $0 < \alpha < 1$, and the transport coefficient Γ is generally analogous to $4Dt^{1-\alpha}$.

The autocorrelation of multiple diffusing species is just a linear combination of the autocorrelations for each species

separately, weighted by their fluorescence intensity (Hess et al., 2002):

$$G_D(\tau) = \sum_{i=1}^{m} \bar{F}_i^2 \cdot G_i(\tau) \Bigg/ \left[\sum_{i=1}^{m} \bar{F}_i^2 \right] \quad (10)$$

where $\bar{F}_i$ and $G_i(\tau)$ are the average fluorescence and autocorrelation for the ith species, respectively, and m is the total number of species. Processes occurring faster than the diffusion time τ_D can be resolved as additional factors that generate shoulders added to $G(\tau)$ at $\tau < \tau_D$. Because time scales are involved that do not depend on the concentration of fluorescent molecules, FCS can be used as a "ratiometric" means of measurement, i.e., it can also extract concentration-independent information from a system.

Consider now transitions of the form B* ↔ B, between bright (B*) and dark (B) states of a molecule or mobile object, which occur on time scales faster than diffusion and are hence observable by FCS, as in triplet-state intersystem crossing, molecular conformational changes, or certain chemical reactions. If such a reaction has an equilibrium constant K, forward and backward reaction rate constants k_+ and k_-, respectively, and a free energy ΔG, as related by

$$K = \frac{[B]}{[B^*]} = \frac{F_B}{[1 - F_B]} = \frac{k_+}{k_-} = \exp(-\Delta G/k_B T) \quad (11)$$

the Boltzmann constant k_B, temperature T, and respective concentrations [B*] and [B], then a characteristic time constant $(\tau_B)^{-1} = k_+ + k_-$, and molecular dark fraction $F_B = [B]/([B^*] + [B])$ may be defined. These quantities τ_B and F_B describe the reaction and are introduced as an exponential factor $F_B \exp(-\tau/\tau_B)$ in the autocorrelation $G(\tau)$, resulting in

$$G(\tau) = G_C G_D$$

$$= \left[\frac{1 - F_B + F_B e^{-\tau/\tau_B}}{1 - F_B} \right] \frac{1}{N}$$

$$\cdot \left[1 + \frac{\tau}{\tau_D} \right]^{-1} \left[1 + \frac{\tau}{\omega^2 \tau_D} \right]^{-0.5} \quad (12)$$

Exploiting the fact that $G_D(0)^{-1} = N$ and $C_0 = N/V$, the physical volume of the observation volume (V) can be determined from an FCS measurement at known concentration C_0, or for a known observation volume an absolute concentration can be obtained.

For multiple independent chemical kinetic processes that occur on well-separated time scales, the fitting function is a product of the diffusion and m chemical kinetic factors (Hess et al., 2002):

$$G_C(\tau) = \prod_{i=1}^{m} [1 - F_i + F_i e^{-\tau/\tau_i}]/[1 - F_i] \quad (13)$$

with amplitude F_i and time scale τ_i.

The above derivation assumes a Gaussian profile for O(**r**). However, if the profile is non-Gaussian, $G_C(\tau)$, $G_D(\tau)$, and $G(\tau)$ will all be distorted, and the kinetic parameters obtained from measurements will be subjected to significant systematic errors. An understanding of the effects of the optics on the functional form of $G(\tau)$ are thus crucial. Therefore, we consider the optical problem of calculation of O(**r**) that depends on the illumination and collection point spread functions of the FCS system, and which allows us to calculate $G(\tau)$ from O(**r**).

Calculation of the point spread function

The point spread function for a lens system is defined as the 3D image of a point source. It is also the intensity distribution in the vicinity of the focal plane resulting from a point source of monochromatic light in the image plane of the lens system. The paraxial approximation (Born and Wolf, 1991) is used to simulate confocal optical systems (Sheppard et al., 1991; Sandison and Webb, 1994), which is adequate for sin $\alpha < 0.8$ (Sandison and Webb, 1994), or roughly NA = 1.06 in water, where $\alpha = \sin^{-1}$ (NA/n) is the half-angle of opening of the objective lens in radians, and n is the refractive index. Here we consider a non-paraxial, high-NA description (Richards and Wolf, 1959), which is a better approximation for $\alpha > 0.8$ and ideal lenses (no aberrations).

The dimensionless optical coordinates used are v and u, corresponding to radial and axial coordinates, respectively:

$$v = kr \sin \alpha \quad (14A)$$

$$u = kz \sin^2 \alpha = \frac{2\pi n \sin^2 \alpha}{\lambda} \cdot z \quad (14B)$$

where k is wavenumber in the medium, λ is wavelength, and $r^2 = x^2 + y^2$. The diffraction theory by Richards and Wolf uses the complex integral representation:

$$\psi_0(u, v) = \int_0^\alpha A(\theta)\sin \theta(1 + \cos \theta)$$

$$\times J_0\left(\frac{v \sin \theta}{\sin \alpha}\right) e^{iu \cos \theta/\sin^2 \alpha} d\theta$$

$$\psi_1(u, v) = \int_0^\alpha A(\theta)\sin^2 \theta J_1\left(\frac{v \sin \theta}{\sin \alpha}\right) e^{iu \cos \theta/\sin^2 \alpha} d\theta \quad (15)$$

$$\psi_2(u, v) = \int_0^\alpha A(\theta)\sin \theta(1 - \cos \theta)$$

$$\times J_2\left(\frac{v \sin \theta}{\sin \alpha}\right) e^{iu \cos \theta/\sin^2 \alpha} d\theta$$

for the energy density (intensity):

$$\langle w_e(u, v, \phi) \rangle = \frac{A_0^2}{16\pi} \{|\psi_0^2| + 4|\psi_1^2|\cos^2 \phi + |\psi_2^2|$$

$$+ 2 \cos 2\phi \, \mathrm{Re}(\psi_0 \psi_2^*)\} \quad (16)$$

where A_0 is a constant, J_i are Bessel functions of ith order, ψ_i are the integrals (Eqs. 15), θ is the integration angle, and ϕ is the azimuthal angle around the longitudinal axis ($\phi = 0$ corresponds to the x-axis).

Illumination optical geometry: the underfilling fraction

Underfilling the back-aperture of the objective can be used to dramatically elongate and enlarge the illumination profile at the focus of the objective and to reduce the effective NA of illumination, which produces a nearly Gaussian illumination profile. Specifically, for a Gaussian beam with $1/e^2$ intensity radius r_0 and objective back-aperture radius r_{BA}, the underfilling fraction $\beta = r_{BA}/r_0$ is just the ratio of the back-aperture radius to the beam radius. For $\beta < 1$ the objective is "overfilled," and for $\beta > 1$ the objective is "underfilled." The apodization function in Eqs. 15

$$A(\theta) = \left[\cos \theta \cdot \exp\left(-2 \frac{r_{BA}^2 \sin^2 \theta}{r_0^2 \sin^2 \alpha} \right) \right]^{0.5}$$

$$= \left[\cos \theta \cdot \exp\left(-2\beta^2 \frac{\sin^2 \theta}{\sin^2 \alpha} \right) \right]^{0.5} \quad (17)$$

(W. Zipfel, unpublished data) is the factor that accounts for the effect of β on the illumination profile.

Collection optical geometry and observation volume profile

After illumination with confocal optics, the second factor determining the optical geometry in FCS is the collection profile, which accounts for the confocal detector aperture and diffraction of collected fluorescence propagating from the object space into the image space, and is combined with the illumination profile to calculate the observation volume profile. The relative spatial collection efficiency can be expressed as the contribution to the measured fluorescence signal as a function of position in object space. Therefore we consider a point $\mathbf{r}_O$ in object space and the contribution to the image it produces (a point spread function centered at the corresponding image point $\mathbf{r}_i$, weighted by the illumination at $\mathbf{r}_O$). The collected portion of the fluorescence emitted from this point gives the collection efficiency for this point relative to other points in the object space. We assume that the detector collects all of the fluorescence that passes through the confocal detector aperture, which is by definition in the image plane.

In the case of two-photon excitation, where there is typically no confocal detector aperture, the observation volume is determined exclusively by the squared illumination intensity:

$$O(\mathbf{r}_O) = S^2(\mathbf{r}_O) \quad (18)$$

where $S(\mathbf{r}_O)$ is the normalized illumination intensity point spread function, which depends on the illumination wavelength, intensity, and objective properties.

In confocal FCS, the collection through the detector aperture must also be considered. The observation volume is defined not only by the illumination profile factor $S(\mathbf{r}_O)$, but also by a collection profile factor $\Omega(\mathbf{r}_O)$ resulting from the objective and tube lens properties, fluorescence emission spectrum, and the confocal detector aperture in the image plane, assuming incoherent emission (Born and Wolf, 1991; Sandison et al., 1995). The observation volume is thus defined as:

$$O(\mathbf{r}_O) = S(\mathbf{r}_O) \int_{\mathrm{Det}} \Omega\left(\frac{\mathbf{r}_i}{M} - \mathbf{r}_O \right) d^2\mathbf{r}_i \quad (19)$$

where $\mathbf{r}_i$ is the position of the element of area in the image plane, $\Omega(\mathbf{r}_O)$ is the centrosymmetric collection point spread function, M is the total magnification of the system, and $\mathbf{r}_O$ is the object-space position. Note that an infinitesimal detector is represented by a delta-function at $\mathbf{r}_i = 0$ and in this case the integral reduces to

$$O(\mathbf{r}_O) = S(\mathbf{r}_O) \int \Omega\left(\frac{\mathbf{r}_i}{M} - \mathbf{r}_O \right) \delta(\mathbf{r}_i = 0) d^2\mathbf{r}_i$$

$$= S(\mathbf{r}_O) \cdot \Omega(-\mathbf{r}_O) = S(\mathbf{r}_O) \cdot \Omega(\mathbf{r}_O). \quad (20)$$

This form reduces to the point spread function of a two-photon (2P) microscope with no detector aperture, where $\Omega(\mathbf{r}_O) \rightarrow S(\mathbf{r}_O)$, yielding $O(\mathbf{r}_O) \rightarrow [S(\mathbf{r}_O)]^2$. The detector aperture size in optical units is calculated using

$$r_d = \frac{2\pi n \sin \alpha}{\lambda \cdot M} \cdot R_d \quad (21)$$

where R_d is the detector aperture radius in real space, and r_d is its radius in dimensionless optical units.

Calculation of collected fluorescence for incoherent emission

The time-averaged collected fluorescence $\langle F \rangle$ is given by an integral over the object space

$$\langle F \rangle = \chi I(\mathbf{r}_O = 0) \cdot \int O(\mathbf{r}_O) C(\mathbf{r}_O) d^3\mathbf{r}_O \quad (22)$$

where $O(\mathbf{r})$ is the observation volume spatial profile, and χ is a constant proportional to overall detection efficiency,

dye fluorescence excitation cross section, and quantum yield.

Observation volume

We define the volume of an illumination or observation profile (note that volume is to be distinguished from profile) by an integral over the appropriate space:

$$V = \left[\int W(\mathbf{r})d\mathbf{r} \right]^2 \times \left[\int W^2(\mathbf{r})d\mathbf{r} \right]^{-1} \qquad (23)$$

where $W(\mathbf{r})$ is the spatial profile normalized to unity at its maximum, i.e., $W(\mathbf{r}) = O(\mathbf{r})/O(\mathbf{r} = 0)$. This definition is similar to established definitions (Thompson, 1991; Mertz et al., 1995; Xu and Webb, 1997). Equation 23 is used to calculate the effective volume of the calculated observation volumes.

Signal-to-noise ratio in FCS

Forming a correlation requires the observation of correlated photons. The S/N ratio in FCS depends simply on being able to observe pairs of photons emitted by a single molecule within a correlation time τ. There must be an abundance of photons within the time window τ for the correlation to have a reasonable S/N ratio, and hence the count rate per molecule ($\eta = F/N$) has been shown to be directly related to signal-to-noise in FCS (Koppel, 1974). In the regime ($\eta\tau \ll 1$), the S/N ratio is proportional to η (Koppel, 1974). The average fluorescence F is measured routinely during FCS, and the zero-time value of the autocorrelation function, $G(0)$, gives the inverse of the mean number of fluorescent molecules in the focal volume (N), if properly corrected for background fluorescence. Corrections to N and η for measured (noncorrelating) background (Koppel, 1974) are made using

$$N = N_{\text{meas}} \frac{\langle F \rangle^2}{\langle F + B \rangle^2} \qquad \eta = \eta_{\text{meas}} \frac{\langle F + B \rangle}{\langle F \rangle}. \qquad (24)$$

Therefore, η is an accessible experimental parameter that is commonly used to optimize the S/N ratio. Fortunately, η is also accessible to our calculations once we have obtained fluorescence and number of molecules (i.e., the physical volume of the observation volume). Because the maximization of S/N with minimal experimental artifacts is of great interest, the dependence of η and artifacts on β and r_d are calculated and examined under the same conditions.

Dependence of count rate per molecule on detector aperture

Small detector aperture limit

The value of η is greatly reduced by using a pinhole smaller than the optimal size. This reduction is dictated by the rapid

decrease in collected fluorescence as the detector gets very small (F is proportional to r_d^2). For small pinhole values, however, the size of the observation volume does not go to zero as the pinhole becomes infinitesimally small; instead, the profile of the observation volume approaches the illumination profile squared. Thus, $F/V \propto \eta \rightarrow 0$ as $r_d \rightarrow 0$.

Large detector aperture limit

When the observation volume is non-Gaussian, the count rate per molecule η also declines for very large r_d. This is clear if one considers η with spatially uniform concentration C_0:

$$\eta = \frac{F}{N} = \frac{F}{C_0 V} = \chi I_0 O \Big|_{\text{ro}=0} \frac{\int W^2(\mathbf{r_O})d\mathbf{r_O}}{\int W(\mathbf{r_O})d\mathbf{r_O}} \qquad (25)$$

where χ is a constant (see above), $I_0 = I(\mathbf{r_O} = 0)$, and $W(\mathbf{r_O}) = O(\mathbf{r_O})/O(\mathbf{r_O} = 0)$. Note that $O(\mathbf{r_O} = 0)$ is proportional to the light collected from a point source at the origin as a function of detector aperture, which reaches a maximum (constant) value at very large aperture. Furthermore, as the pinhole is opened, the observation profile more and more closely approaches the illumination profile. In the presence of diffraction fringes in the observation volume at large r and z, the integral of W will grow more quickly than the integral of W^2, which contains the profile to the second power, which greatly reduces the influence of the dim regions. Hence, at large detector apertures, the product of $O|_{r=0} \cdot \int W^2(\mathbf{r})d\mathbf{r}/ \int W(\mathbf{r})d\mathbf{r}$ declines, resulting in decreased count rate per molecule. The best detector aperture (to give the maximum η) is therefore a compromise between reduced collected fluorescence for small aperture values and large observation volumes, which occur at large aperture values.

No peak in $\eta(r_d)$ with a 3D-ellipsoidal Gaussian focal volume

If a 3D-ellipsoidal Gaussian profile is assumed for illumination and collection functions, as well as a Gaussian profile for convolution with the pinhole, then the observation volume is Gaussian. With a Gaussian observation volume and spatially constant concentration C_0 the ratio $\int W^2(\mathbf{r})d\mathbf{r}/ \int W(\mathbf{r})d\mathbf{r}$ is constant, and hence the count rate per molecule $\eta(r_d)$ is proportional simply to $O(\mathbf{r_O} = 0)$, which is proportional to the collected fluorescence from a point source at the origin (object plane focus) and will only increase with increased detector aperture. Therefore, using a Gaussian approximation for the observation volume will never give a peak in count rate per molecule as a function of detector aperture.

MATERIALS AND METHODS

Numerical implementation

The integrations in Eqs. 15 are done numerically using Bessel functions $J_0(x)$, $J_1(x)$, and $J_2(x)$ evaluated from $x = 0$ to 200 in steps of 0.02 and interpolated linearly. Illumination profiles agree well at low NA with paraxial results in the literature (data not shown) and with paraxial calculations (Sandison and Webb, 1994) at intermediate NA. Furthermore, results as a function of β at high NA agree with independent calculations by the method of Richards and Wolf and measurements (W. Zipfel, personal communication, 2001). Numerically, the observation volume profile (Eq. 19) is calculated as a sum:

$$O(\mathbf{r}_O) = \sum_{x=-n\Delta x}^{x=n\Delta x} \sum_{x=-n\Delta y}^{y=n\Delta y} S(\mathbf{r}_O) \cdot \Omega(\mathbf{r} - \mathbf{r}_O)\Theta(\mathbf{r}_d - \mathbf{r})\Delta x \Delta y$$

(26)

where $\mathbf{r} = (x, y, 0)$ and n is the number of steps in the grid approximating the detector face, which is contained in the x-y plane. The theta function excludes values of $\mathbf{r}$ outside the detector radius. To accelerate calculations, the values for the normalized illumination intensity $S(u, v)$ and collection efficiency $\Omega(u, v)$ are calculated once and stored in 1000×1000 element floating-point arrays. Values are then interpolated from the array values for use in the observation volume calculations. Typical parameters used are $\alpha = 1.12$ ($=1.2$ NA in water), v from 0 to 200, u from 0 to 200, 1000 steps in both u and v, $\phi = 0$, $\Delta\theta = 0.002$, and $\beta \ll 1$ for overfilled or $\beta > 1$ for underfilled. Typically a 20×20 detector grid is used.

Calculated correlation curves

The double integral in Eq. 6 is solved numerically as a convolution, followed by a product, followed by a single integral in three dimensions, using a discrete three-dimensional array of $(128)^3 = 2,097,152$ elements. The theoretical $G(\tau)$ in Eq. 6 then becomes (for discrete elements)

$$G_T(\tau) = \sum_{i,j,k} T_{ijk}(\tau)$$

(27)

where T_{ijk} is a three-dimensional array representing the result of the first spatial integral, calculated from a product of the discretized spatial observation profile O_{ijk}, and the result of the convolution, A_{ijk}. The subscripts i, j, k denote spatial position along the three Cartesian axes. The convolution is between the concentration function C_{ijk} and a second (identical) spatial observation profile, O'_{ijk}. C_{ijk} is the solution to the diffusion equation from Eq. 5 evaluated at the discrete grid positions defined by the three orthogonal Cartesian coordinates i, j, and k. Note that the Einstein summation convention is *not* used below.

$$T_{ijk}(\tau) = O_{ijk}A_{ijk}$$

(28)

$$A_{ijk} = \text{FFT}^{-1}(\tilde{C}_{ijk}\tilde{O}'_{ijk})$$

(29)

where

$$\tilde{C}_{ijk} = \tilde{C}_{ijk}(\tau) = \text{FFT}(C_{ijk}(\tau)).$$

(30)

The theoretical $G_T(\tau)$ is calculated at 50 logarithmically spaced τ values spanning $\sim$4 orders of magnitude in time. $G_T(\tau)$ can also be used to obtain a diffusion coefficient by fitting $G_T(\tau)$ to the measured $G(\tau)$, using the diffusion coefficient as a fitting parameter. The theoretical $G_T(\tau)$ is normalized such that $G(\tau \to 0) = 1$. The two-photon autocorrelation was calculated using the square of the illumination profile for O_{ijk} and O'_{ijk}, and using the same C_{ijk} as in the one-photon case. Because there is usually no detector aperture in two-photon FCS, the illumination profile is the only factor determining the profile of the observation volume.

Experimental methods

Autocorrelation curves were measured with two different systems.

One-photon FCS with aperture

A Confocor (Carl Zeiss, Jena, Germany) FCS instrument used the 488 nm line of a 20 mW argon ion laser (Zeiss), focused by a Zeiss C-Apochromat infinity-corrected 1.2 NA 40× water objective ($\beta \sim 1$) and either the standard 8-well plastic sample holders designed for use with the system or a deep-well slide covered with a no. 1½ coverslip. The fluorophore was rhodamine green (D-6107, Molecular Probes, Eugene, OR) of known concentration between 1 and 300 nM. Fluorescence was collected through the same objective and directed through FITC filters (480/10 excitation filter, 510 LP dichroic, 540/40 emission filter) and a variable-diameter aperture (25–200 μm) onto an avalanche photodiode (EG&G, Salem, MA, provided by Zeiss). The measured overall magnification was 81. The fluorescence signal is digitized and autocorrelation curves thereby calculated by a PC with onboard correlator card (ALV Laser, Hamburg, Germany). Software provided with the Confocor was used to do numerical fits to the data and obtain diffusion time (τ_d), number of molecules N, average intensity $\langle F \rangle$, and count rate per molecule η. Multiple correlation curves were obtained for each detector aperture r_d on different days. Measured parameters were corrected for measured background from a distilled water blank under the same conditions. Unattenuated laser power at the back-aperture of the objective was 12 mW, but the intensity was always further reduced by neutral density filters to typically 10–100 μW. Measurements of the number of molecules were done at the highest concentrations ($\sim$100–300 nM) to eliminate dye-glass sticking artifacts and to enable measurement of the absorbance of the sample accurately.

Two-photon FCS

A $\sim$100 fs pulsed-IR Ti:sapphire laser (Tsunami, Spectra-Physics, Palo Alto, CA) pumped by an 8 W argon ion laser (Beamlok, Spectra-Physics) mode-locked at 980 nm was guided into an IM35 microscope (Zeiss, Jena, Germany) and reflected by an appropriate dichroic into the back-aperture of a 1.2 NA 60× objective (Zeiss). Fluorescence was passed by the same dichroic and focused by the tube lens into a multimode optical fiber (200 μm $\sim$ 52 ou nominal diameter). The fiber carried the fluorescence to the detector, an avalanche photodiode (model SPCM-AQ-141-FC, EG&G) whose active area was large enough not to restrict the observation volume in the absence of light scattering (Schwille et al., 1999a). The TTL pulses from the detector are autocorrelated in a PC by an autocorrelator card (ALV/5000, ALV Laser). Typically, 20 runs of 30–60 s each were averaged. FCS on rhodamine green or Alexa488 (Molecular Probes) at $\sim$10 nM concentration in a deep-well slide with a no. 1½ coverslip was performed for 20 scans of 60 s each, per intensity. Power at the sample was 5–10 mW, while the back-aperture of the objective was underfilled ($\beta \sim 3$) or overfilled ($\beta < 1$). One-photon FCS results using the second setup, 488 nm illumination from an argon ion laser, and various fiber diameters agreed with the results from the first setup.

The concentration of rhodamine green was measured by absorbance in a spectrophotometer (HP 8451A, Hewlett Packard, Palo Alto, CA) in either a 3×3 mm or 10×10 mm quartz cuvette. Between 5 and 15 independent wavelength scans were averaged. We used the peak absorbance (at $\lambda \sim 490$ nm) subtracting the baseline at long wavelength ($\lambda \sim 650$ nm) in the equation $C = A/\varepsilon L$ (with $\varepsilon = 5.8 \times 10^4$ M^{-1} cm^{-1}, A = absorbance, $L = 0.3$ or 1.0 cm) to calculate concentration.

RESULTS AND DISCUSSION

Calculated illumination and observation volume profiles are non-Gaussian and will lead to artifacts

We delineate the effects of the observation volume profile $O(\mathbf{r}_O)$ on the FCS autocorrelation function and the resulting

consequences on the measured dynamics and analytical FCS parameters. Here we will present calculated illumination and observation profiles for a high-NA objective, as a function of the back-aperture underfilling (β) and the confocal detector aperture (for observation profiles only), first as two-dimensional pseudocolor intensity bitmaps, then as plots with attempted fits to analytic functions. It will be shown that the observation profile is predicted to be non-Gaussian under many experimental conditions. The detector-aperture dependence of the count rate per molecule, which is directly related to the signal-to-noise ratio, provides direct evidence for the non-Gaussian nature of the observation volume.

Next, the effect of a non-Gaussian observation volume on the FCS autocorrelation function is presented, demonstrating that inaccuracy or artifacts frequently arise as a result of using the standard fitting function (i.e., Eq. 7). The form and magnitude of these artifacts are then explored; their effects on the measured diffusion time, diffusion coefficient, axial ratio, and exponential fraction(s) and time constant(s) are shown. Methods for avoiding artifacts are outlined. Finally, the collected fluorescence and volume of observation (number of molecules at a known concentration) are analyzed as a function of β and r_d and provide evidence for an actual observation volume that is larger than predicted.

Calculated illumination and observation volume profiles

It is known that the intensity profile at the focus of an underfilled, thin, low-NA lens is a Gaussian in the radial direction and a Lorentzian in the axial direction (Self, 1983). FCS typically uses a high-NA objective lens for intense illumination and efficient fluorescence collection, and the illumination and observation profiles are typically assumed to be 3D-ellipsoidal Gaussians. However, it will be shown that this assumption is not sufficiently accurate for precise analysis under many common measurement conditions.

Fig. 1 visually depicts FCS observation profiles as logarithmically scaled pseudocolor bitmaps, with various optical parameters. Here one can readily see evidence for the non-Gaussian profiles and recognize the dependence of the observation volume on the key experimental variables β, objective underfilling fraction, and r_d, confocal detector aperture radius. Most importantly, the observation profile for an overfilled back-aperture ($\beta \ll 1$), commonly used in FCS (*top row*), has visible oscillatory aperture-limited diffraction fringes, especially at large r_d, where the observation profile is unconstrained and approximately equal to the illumination profile (*left side*). These fringes clearly indicate non-Gaussian profiles, as a Gaussian would decay monotonically to zero without oscillation. Second, closing the detector aperture (moving from right to left in the figure) constrains the volume significantly, reduces fluorescence collection from regions far from the focus (i.e., the fringes),

and as will be shown, results in a more nearly Gaussian observation profile. Third, when the objective is underfilled ($\beta = 1.25$; *second row*) and severely underfilled ($\beta = 4$; *third row*), the illumination profile and observation volume are elongated in the axial direction, and somewhat in the radial direction, but this effect is significant only for $\beta > 1.25$. Underfilling also reduces or eliminates the diffraction fringes and results in a smoother, more nearly Gaussian observation volume (shown in the *bottom row*).

To be more quantitative, we now move from the images to plots of the corresponding profiles of the illumination and observation volume along the axial and radial directions, which will show functional dependence. First we consider the functional form of the axial illumination profiles, which corresponds approximately to the limiting case of the observation profile with very large r_d and overfilled back-aperture (*top left image* of Fig. 1). Figs. 2 and 3 show axial and radial plots, respectively, of the illumination profile for an overfilled 1.2 NA objective, which indicates that the functional form is neither Gaussian nor Lorentzian in the axial direction. Plotting $\{-\text{Ln}[O(u, 0)/O(0, 0)]\}^{0.5}$ emphasizes the deviation from Gaussian because an actual Gaussian, $A \exp(-Bu^2)$, plotted this way is $\{-\text{Ln}[Ae^{-Bu^2}/A]\}^{0.5} = \{Bu^2\}^{0.5} = u(B^{0.5})$ which is linear in u. In general, the Gaussian fits well at short distances from the focus, but decays too quickly at large distances. The deviation begins at distances near or slightly less than the $1/e^2$ waist and worsens at larger distances. The overfilled radial illumination profile is also neither a Lorentzian nor a Gaussian beyond the $1/e^2$ waist (see oscillations in Fig. 1, *top left*).

Because the observation volume is a function of the illumination profile (Eq. 3), it is not surprising that the observation profile is also non-Gaussian, as it is shown plotted with attempted Gaussian fits along the radial direction (see Fig. 4, *left*) and axial direction (see Fig. 4, *right*). Again, the Gaussian is only a reasonable approximation very close to the focus. Furthermore, because the fringes are at such large distances from the focus, they occupy a significant volume due to the approximate axial symmetry of the system, and can have a strong influence on the measured volume and autocorrelation in FCS. Therefore, in many circumstances, no simple analytic function will be able to describe the FCS observation volume accurately.

An experimental quantity derived directly from the calculated observation volumes is the count rate per molecule (η), which is directly related to the S/N in FCS. While a typical FCS fit using N and τ_D as free parameters does give information about the observation volume, in the absence of fluorescence saturation and photobleaching the aperture-dependence of η provides a direct experimental test of whether $O(\mathbf{r}_O)$ is Gaussian (see also Eq. 25), because a Gaussian observation volume will not produce a peak in η as a function of detector aperture r_d. Hence, important physical properties of the observation volume can be extracted from the dependence of η on r_d.
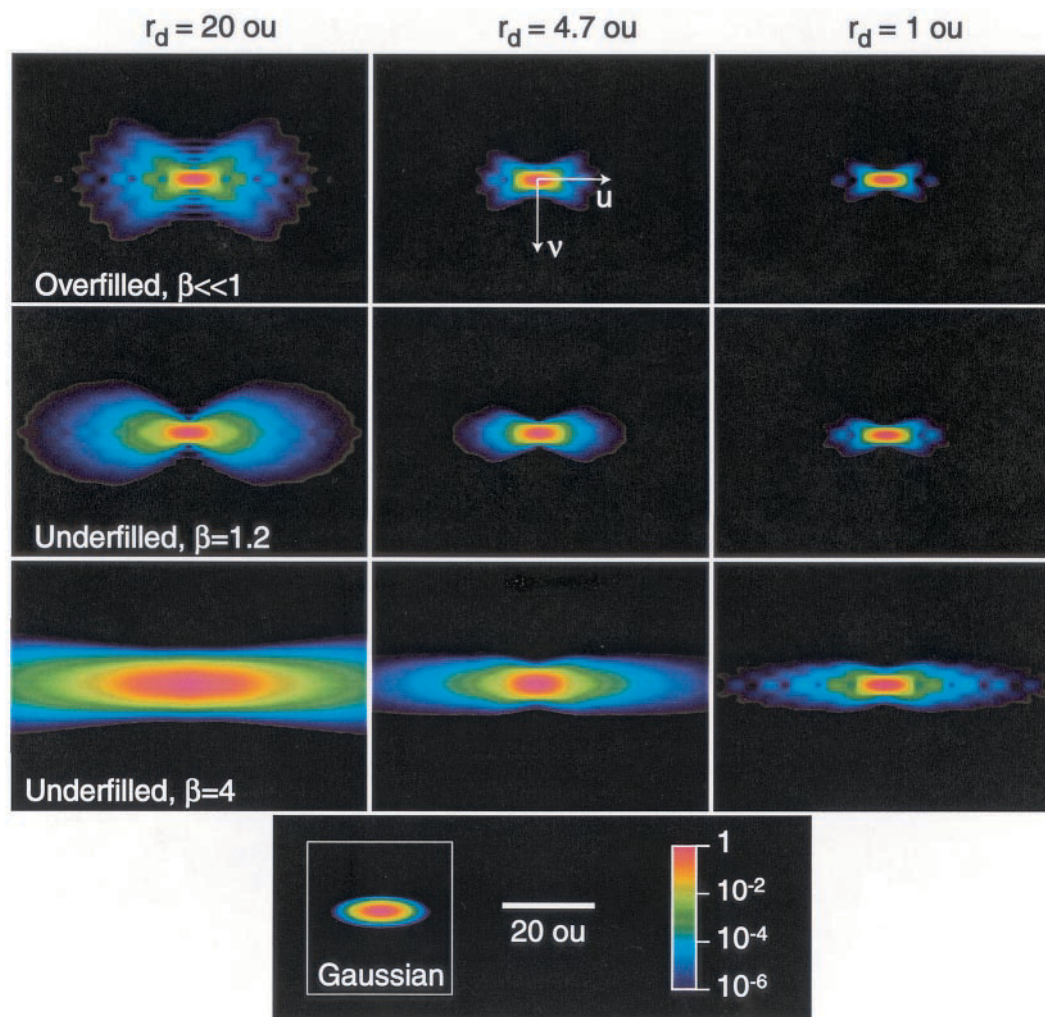
FIGURE 1   Observation profile images in log-scale pseudocolor, for a 1.2 NA objective, plotted for various values of the detector aperture r_d and underfilling fraction β. Images show the axial direction (u) horizontally and the radial direction (v) vertically away from the focus. The logarithmic color scale emphasizes the low-amplitude fringes away from the focus. Closing the detector aperture (*from left to right*) constrains the observation volume, reducing fluorescence collection from the tails of the illumination profile that are particularly periodic and non-Gaussian for an overfilled back aperture (*top row*). Underfilling the back aperture (*from top to bottom*) elongates and smoothes the illumination profile and hence the observation volume, but the effect is not drastic until $\beta > 1.25$. Using small r_d and an underfilled back aperture ($\beta > 1$) results in a more nearly Gaussian observation volume. An exactly Gaussian profile is shown (*bottom row left*) with a 20 ou scale bar (*center*) and intensity color scale (*right*) for reference.

Measurements confirm observation volume is non-Gaussian

Peak in count rate per molecule versus detector aperture

An important result is the dependence of the count rate per molecule (η) on confocal detector aperture size (r_d). A peak in η is observed, which demonstrates that 1) there is an optimal (maximal S/N) set of conditions for FCS, and 2) that the observation volume must be non-Gaussian.

Fig. 5 shows the strong maxima of the calculated and measured $\eta(r_d)$, with diffraction theoretical curves shown for overfilled and underfilled back-aperture and a theory curve assuming a Gaussian observation volume. The curves are normalized such that their peak value is unity. The best fit of the aperture dependence is for the overfilled back-aperture, and the peak position does not change significantly for $\beta < 1.6$. The peak does shift to slightly larger values when the objective is underfilled, because as the focal volume is made larger, the pinhole that restricts the observation volume enough to optimize η will also be larger.

Previous studies on confocal microscopy also discuss optimization of signal to noise (Webb et al., 1990; Sandison and Webb, 1994; Sandison et al., 1995). However, signal-to-noise optimization in confocal FCS is different than in confocal microscopy because FCS has a different optimization criterion. The best detector aperture for FCS is different from the best aperture size for maximal S/N in a confocal
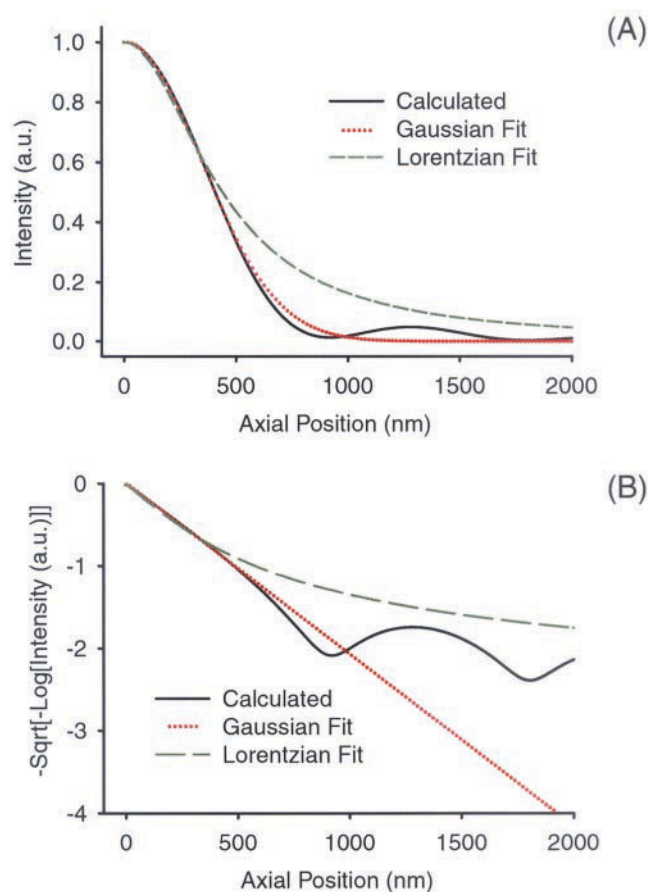
FIGURE 2 Using conventional analytical functions to fit calculated axial illumination profiles is ineffective. The calculated profile for an overfilled 1.2 NA water objective (*solid line*) is fit with a Lorentzian (*dashed line*) and Gaussian (*dotted line*) on a linear (*A*) and logarithmic (*B*) intensity scale.
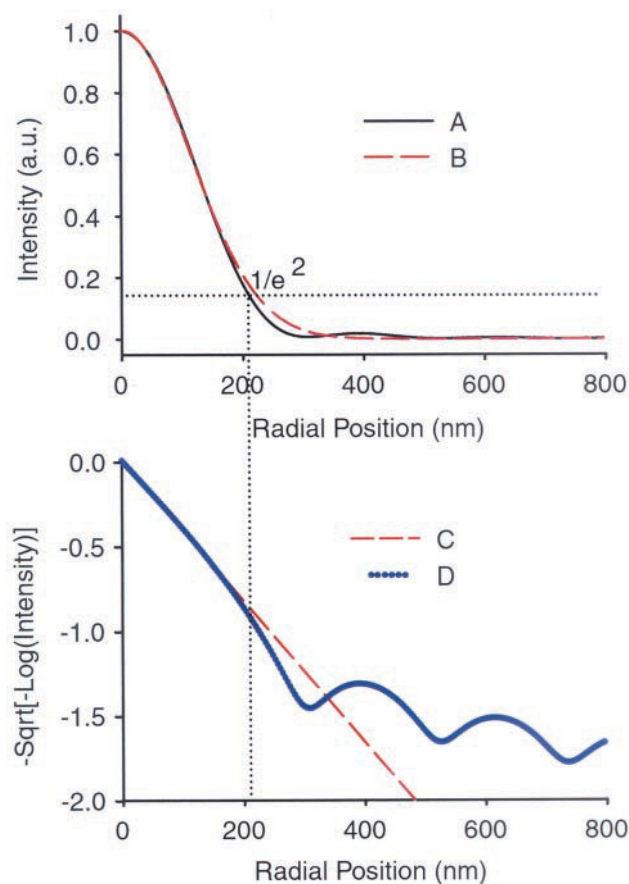
FIGURE 3 Deviation of calculated radial illumination profile from a Gaussian profile. The calculated illumination profile for an overfilled 1.2 NA water objective is shown along the radial axis (*A* and *D*), with a linear intensity scale (*top*) and a nonlinear scale (*bottom*), $[-\log(\text{intensity})]^{0.5}$, which should yield a straight line for all positions if the profile is Gaussian. Note that the profile begins to deviate from Gaussian (*B* and *C*, *red dashed line*) at less than the $1/e^2$ radial width (*dotted black line*).

microscope (using a paraxial diffraction theory) of 2.4–3.3 ou for an overfilled objective (Sandison and Webb, 1994). This difference is expected considering that the above S/N optimization for FCS does not explicitly consider the effects of fluorescence background, and is therefore defined differently. A 3D-ellipsoidal Gaussian observation volume predicts no peak in $\eta(r_d)$; however, a peak is observed in both our experimental and diffraction theory results, evidence that a Gaussian observation volume is not consistent with the measured results.

The peak in count rate per molecule (at $r_d \sim 4.5$ ou for a 40×1.2 NA water immersion objective) also signifies that there is, under many typical measurement conditions, an optimal detector aperture that maximizes η, and hence S/N. Underfilling reduces the maximum η by a factor of ~2 for $\beta = 2.2$, but does not change the location of the peak significantly for $\beta < 1.6$. However, as will be shown, using an overfilled back-aperture and the detector aperture that gives optimal S/N will result in a non-Gaussian observation volume and artifacts in the autocorrelation. Therefore, great

care must be taken by the FCS user who requires the smallest possible observation volume, optimal S/N, and a Gaussian observation profile.

Underfilling the objective decreases the peak value of the count rate per molecule. Fig. 6 shows calculated η versus detector aperture for different underfilling fractions, this time keeping the integrated rate of excitation in the *x-y* plane constant, i.e., constant illumination power. Note the greatly reduced magnitude of η with underfilling, due to decreased intensity at the focus. The effect of underfilling becomes particularly pronounced for $\beta > 1.6$. The peak becomes less pronounced, indicating that the volume is more nearly Gaussian for an underfilled back-aperture. The effect of r_d on η is negligible for $r_d > 7$ ou when $\beta > 4$.

The detector-aperture-dependence of η shows clear experimental evidence of a non-Gaussian observation volume (see Eq. 25). A non-Gaussian illumination profile such as shown in Fig. 1 with an overfilled detector aperture will
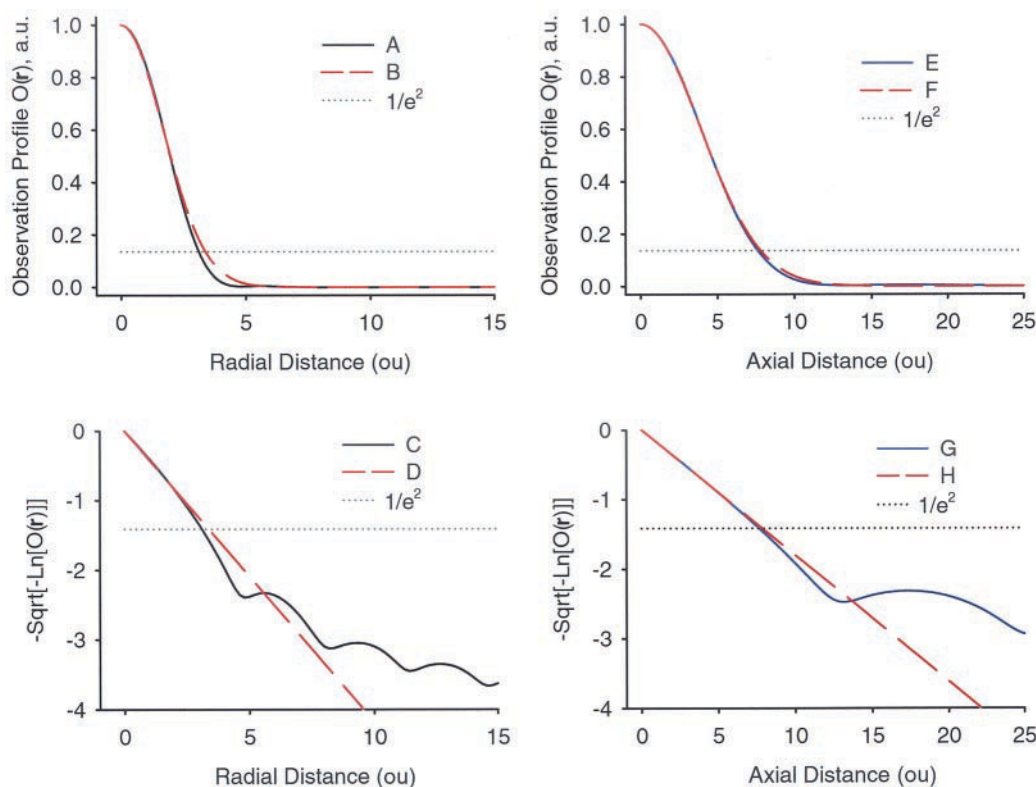
FIGURE 4   Attempt to fit calculated observation profiles for an overfilled 1.2 NA water objective with a 4.7 ou detector aperture (corresponding to the top center panel in Fig. 1) using analytical functions. *Left*: The radial profile $O(\mathbf{r}) = O(u, v) = O(0, v)$ is shown on a linear scale (*A*) with Gaussian fit (*B*), and on a nonlinear scale (*C*) with the same fit (*D*). *Right*: The axial profile $O(\mathbf{r}) = O(u, v) = O(u, 0)$, (*top right, E*) and its Gaussian fit (*F*), is also shown with nonlinear vertical axis (*G*) and fit (*H*). The nonlinear axis transforms any Gaussian function into a straight line.

have fringes far from the focus, which contribute significantly to the volume when the detector aperture is large, increasing the number of weakly fluorescent molecules and reducing the average fluorescence per molecule and the value of η. The peak in η is most pronounced for an overfilled back-aperture where the observation volume is
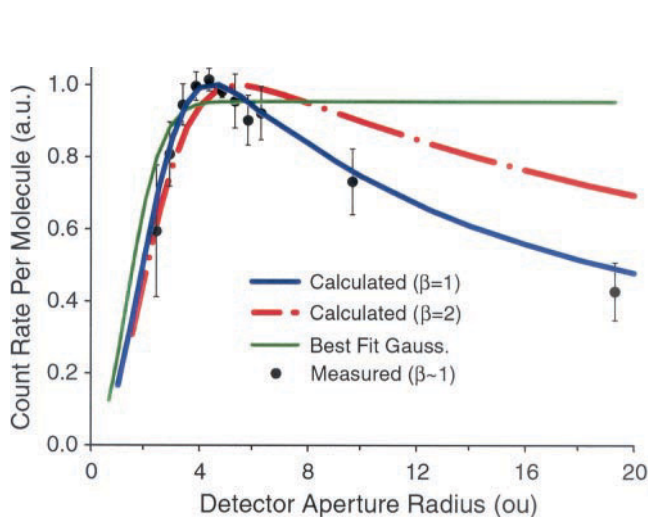

FIGURE 5   Experimental evidence for a non-Gaussian observation volume. Measurements and diffraction-theory-calculated count rate per molecule (η) versus detector aperture (r_d) for a 1.2 NA 40× water objective as a function of underfilling fraction (β). A Gaussian observation profile predicts no maximum in $\eta(r_d)$ at finite r_d (*thin green curve*).
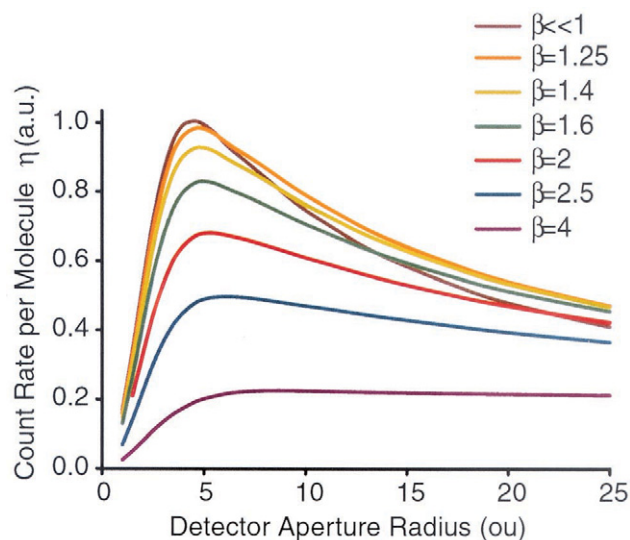

FIGURE 6   Underfilling reduces count rate per molecule. The effect of underfilling fraction (β) and detector radius r_d on count rate per molecule (η) for a 1.2 NA 40× water objective. The curves are normalized such that the total rate of excitations in the *x-y* (focal) plane is constant.
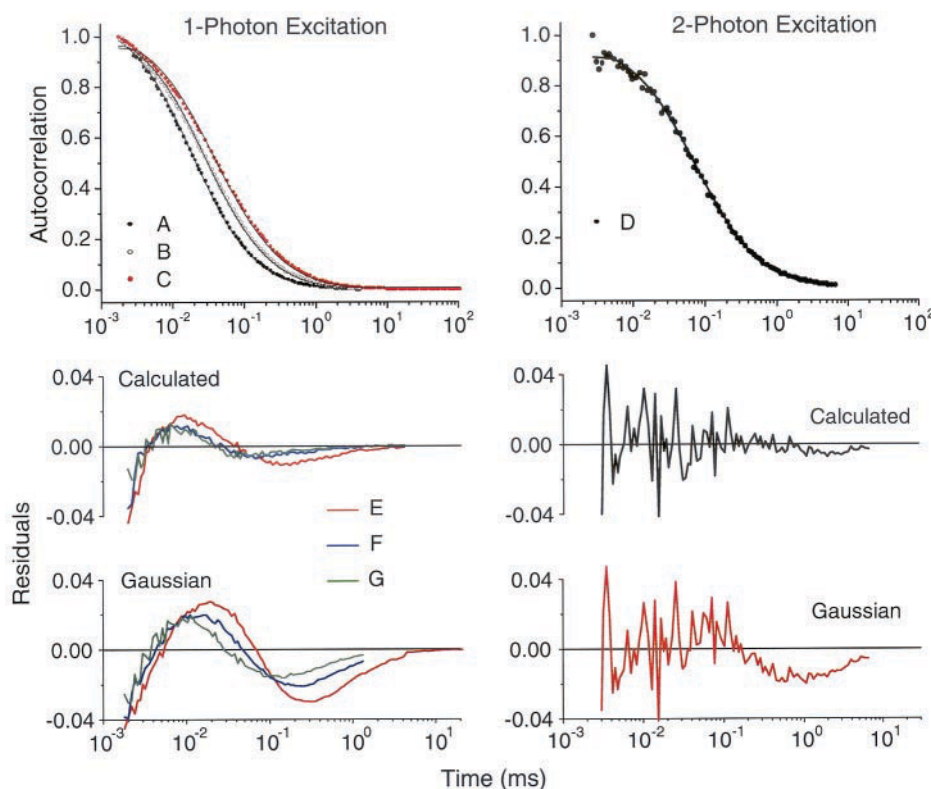
FIGURE 7 Comparison of measured and calculated autocorrelation versus detector aperture. (*Top*) Measured autocorrelation (*points*) and diffraction theory fits (*lines*) for 1P-FCS (*left column*) with detector aperture diameter (*A*) 2.4 ou, (*B*) 3.9 ou, (*C*) 5.8 ou, and for 2P-FCS (*right column*) with no detector aperture (*D*). Residuals are shown below for fits with calculated autocorrelation (*middle row*) and with Eq. 7 (Gaussian volume; *bottom row*). Residual curves (*E–G*) correspond to measured data in *A–C*, respectively. Residuals for 2P-FCS are shown on the middle and bottom right plots. The number of molecules and diffusion coefficient were used as free parameters for the theoretical curve-fitting.

most non-Gaussian due to the diffraction fringes. At larger values of β, where the objective is underfilled, the peak is less pronounced or nonexistent (see Fig. 6). Because the measured FCS diffusion autocorrelation is a function of the observation profile (Eq. 6), deviations from Gaussian behavior in the observation profile will result in deviations from the analytic form for the diffusion autocorrelation (Eq. 7), which assumes Gaussian behavior. Consequently, analysis of the effects of these deviations on the autocorrelation was performed by simulation of the autocorrelation function using calculated non-Gaussian observation profiles (see Fig. 7). Simulated autocorrelation functions also offer a means to measure absolute values of the diffusion coefficient of a molecule.

Simulation of the autocorrelation function using diffraction-based observation profiles

Comparison of calculations with measured autocorrelation and diffusion coefficients

This section demonstrates the degree of agreement between calculations using the predictions of diffraction theory and measured FCS results, including 1) autocorrelation func-

tions at high NA and 2) diffusion coefficients, which are a measure of similarity of the experimental and calculated observation volumes. The measured autocorrelation of rhodamine green (aperture setup) is fit using the calculated autocorrelation for a 1.2 NA 40× water objective, excitation wavelength $\lambda_x = 488$ nm. Fig. 7 shows the measured autocorrelation, fit using the diffraction-calculated autocorrelation, using diffusion coefficient, number of molecules, and a baseline as free parameters.

The diffusion coefficient that produced the best fit against the measured autocorrelation is reported in Fig. 8 for rhodamine green under 1P excitation for a variety of circular detector apertures (average of three measurements each) and under 2P excitation with no detector aperture. FCS under 2P excitation is expected to yield highly reliable values for D because the observation volume is very nearly Gaussian. Under 1P excitation, some systematic variation of D with r_d is observed and is likely due to the observation volume not being adequately predicted by the diffraction theory. A deviation of 10% between actual and predicted waists of the observation volume will result in a $\sim$20% deviation between actual and predicted diffusion coefficients (i.e., 3.0×10^{-6} would become 3.6×10^{-6}) because
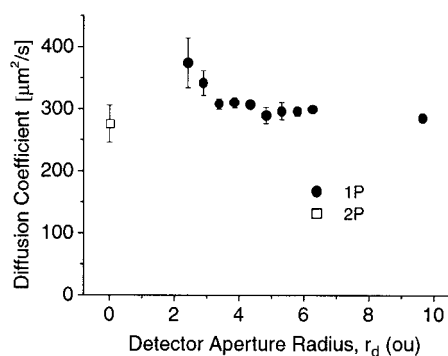
FIGURE 8 Diffusion coefficient of rhodamine green (RG) using diffraction-calculated autocorrelation to fit the measured autocorrelation for the aperture setup (*filled circles*) for various detector aperture radii (r_d). Diffusion coefficient of RG under 2P-excitation (no detector aperture) using a similar procedure (*open square*).

$D \sim \rho^2/4\tau_D$. The average value for the diffusion coefficient over the detector aperture range where the values are approximately constant ($r_d = 3.4$–6.3 ou) is $D = (3.01 \pm 0.08) \times 10^{-6}$ cm^2/s.

The literature values reported (by FCS measurements assuming a Gaussian observation volume) for the diffusion coefficient of rhodamine green $D = 2.33 \times 10^{-6}$ cm^2/s (Visser et al., 1999) and rhodamine 6-G, 2.8×10^{-6} cm^2/s (Meseth et al., 1999; Rigler and Widengren, 1990; Rigler et al., 1993), as well as 3×10^{-6} cm^2/s (Thompson, 1991) are in reasonable agreement with our measurements except for the value from Visser et al. However, it is important to consider the basis for the literature values of D for rhodamine that seem to be reported consistently by 1P-FCS measurements, all of which assume a Gaussian observation volume and, as we will show, may suffer from artifacts. Forcing the measured autocorrelation to be fit using an erroneous value of D may hide systematic kinetics errors or other artifacts, as will be shown below.

Comparison of calculations with analytic fitting function from literature

To directly address the question of whether the assumption of a Gaussian observation volume results in an inferior description of the measured $G_D(\tau)$, comparison is made between the standard fitting function (Eq. 7), which assumes a Gaussian observation volume, and the calculated $G_D(\tau)$, which does not. Fig. 7 also shows fitting residuals to emphasize the small but significant differences between using Eq. 7 and our calculated nonanalytic autocorrelation function using a diffraction-modeled observation volume. The standard fitting function, Eq. 7, using N and τ_D as fitting parameters and $\omega = 2.5$, does not describe the measured autocorrelation as well as the calculation, based on the mean-squared residuals integrated over the time range, which were ~4-fold larger for the Gaussian-volume

fits, indicating that the non-Gaussian diffraction-based observation volume is a better description than a Gaussian one. However, both fitting functions have systematic (i.e., nonrandom) residuals. These systematic residuals are a direct consequence of the effect of the observation volume on the diffusion autocorrelation function. A nonstandard (non-Gaussian) observation volume will result in a nonstandard measured diffusion autocorrelation, which will not be described by Eq. 7. The measured 2P-FCS autocorrelation is described well by the diffraction-calculated autocorrelation and shows minimal systematic deviation. The mean-squared residuals for the Gaussian fitting function (Eq. 7) were only ~30% larger, indicating that the 2P-FCS observation volume is reasonably Gaussian.

The presence of some residual, even using the diffraction-calculated 1P-autocorrelation, implies that the experimental observation volume is also not identical to the predicted observation volume. The fitting residuals are worse for large detector apertures, indicating that the observation volume profile is less well-described under these conditions, possibly due to an illumination profile that is larger than expected because of optical effects not considered in the model. However, the residuals are significantly better for the diffraction theory than for the Gaussian fitting function. Therefore, one can use the calculations to predict systematically the conditions (i.e., detector aperture and objective underfilling fraction) where a non-Gaussian observation volume will occur. The degree of disagreement between the diffraction-calculated autocorrelation and the Gaussian-volume fitting function (Eq. 7) indicates the extent to which the observation volume is non-Gaussian.

Comparison between diffraction-calculated and Gaussian-volume autocorrelation functions: altered form of $G(\tau)$ when observation volume is non-Gaussian can cause artifacts

A nonstandard (non-Gaussian) observation volume will result in a nonstandard measured diffusion autocorrelation function that cannot be fit using the standard analytical FCS fitting function (Eq. 7). To analytically determine the conditions where $G(\tau)$ has a significantly altered form, the diffraction-calculated autocorrelation is fitted using Eq. 7 under various conditions.

Fig. 9 shows calculated autocorrelations that would be measured at low intensity on a system with a 1.2 NA 40× objective and various β values and detector apertures, with attempted fits using Eq. 7. Note that the small detector aperture results in an excellent fit of the calculated autocorrelation using Eq. 7, evidence that even with an overfilled back-aperture, a very small detector aperture will result in a nearly Gaussian observation volume. However, with large detector aperture, the fit of the calculation with Eq. 7 is poor, and therefore the observation volume is non-Gaussian even when the objective is underfilled. Optical fibers used
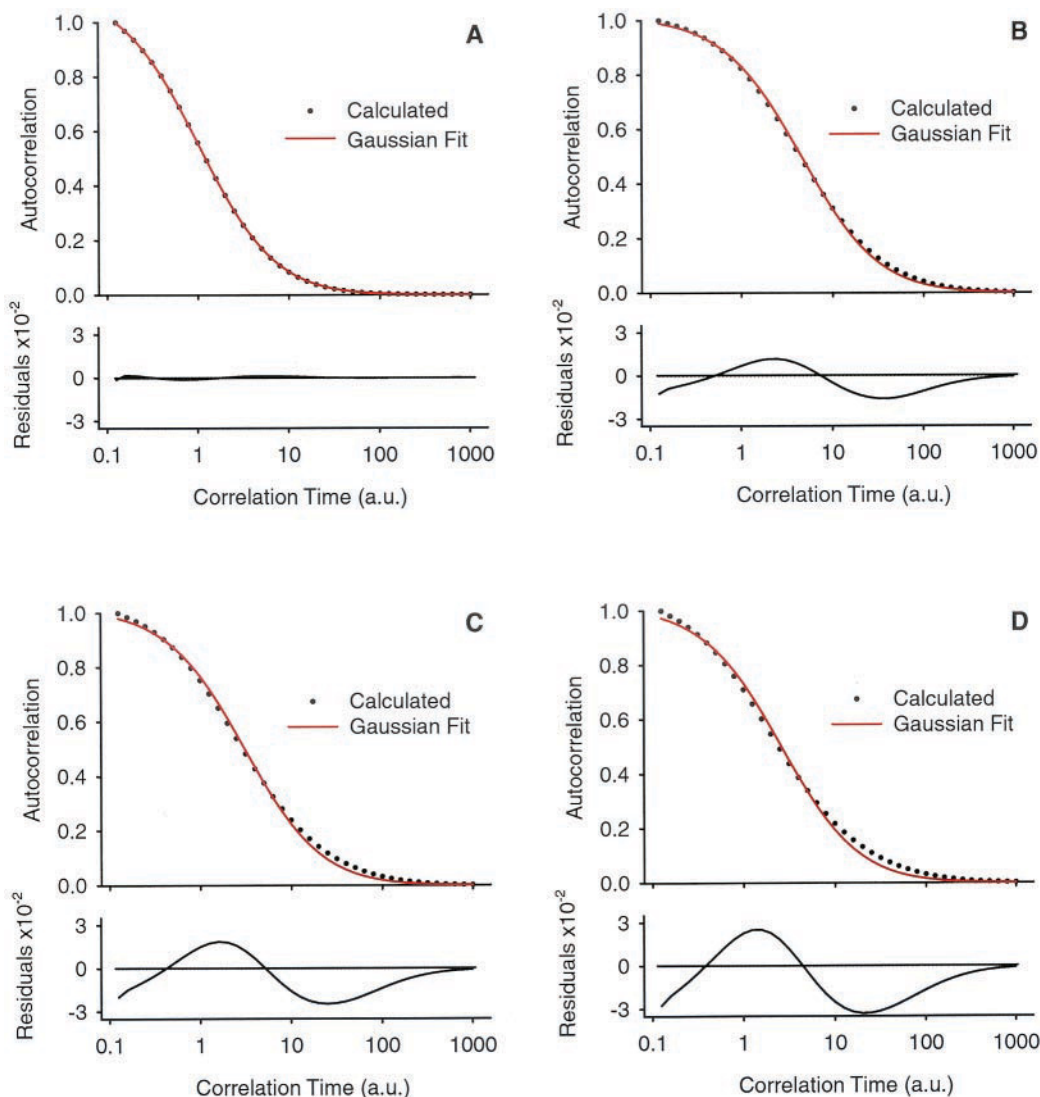
FIGURE 9   Alteration of the autocorrelation by the non-Gaussian collection volume. The calculated autocorrelation (*points*) for a 1.2 NA water objective, using the observation volume modeled by diffraction theory, is compared with an attempted fit (*red line*) using a Gaussian observation volume, which corresponds to the usual analytical fitting function from literature, $G_G(\tau) = N^{-1}[1 + \tau/\tau_D]^{-1}[1 + \tau/\omega^2\tau_D]^{-0.5}$. The fitting residuals are shown directly below (*black line*). (A) $r_d = 1.5$ ou, overfilled ($\beta \ll 1$). (B) $r_d = 20$ ou, underfilled ($\beta = 2$). (C) $r_d = 20$ ou, $\beta = 1.25$. (D) $r_d = 20$ ou, $\beta \ll 1$. The Gaussian profile fits only the case with the nearly closed detector aperture.

as confocal apertures often have larger effective diameters than their stated nominal values (D. Larson, personal communication, 2002) and do not ensure a Gaussian observation volume. The observation volume becomes progressively more non-Gaussian as the objective back-aperture is overfilled, thus preventing the use of the standard function (Eq. 7) for fitting. Therefore, when the observation volume is non-Gaussian (especially when an overfilled back-aperture and large r_d are used), the unmodified standard fitting function cannot be used because it assumes a Gaussian volume.

A non-Gaussian observation volume can drastically affect measured FCS parameters such as the axial ratio (ω). A frequent experimental observation is that the value of ω is larger than expected, often diverging to infinity, which does not represent the physical situation of a confined observation volume with finite waists in both ρ and z. However, in the presence of artifacts resulting from a non-Gaussian observation volume, even weak artifacts such as an amplitude of a few percent (i.e., $F_B \sim 0.05$ fitting with Eq. 12), the intensity fringes far from the focus will contribute to the measured diffusion autocorrelation with significant amplitude due to the relatively large volume they occupy in space. These fringes result in additional correlation amplitude at long time scales, where an increased value of the axial ratio is typically used to improve the fit (see Fig. 10). When
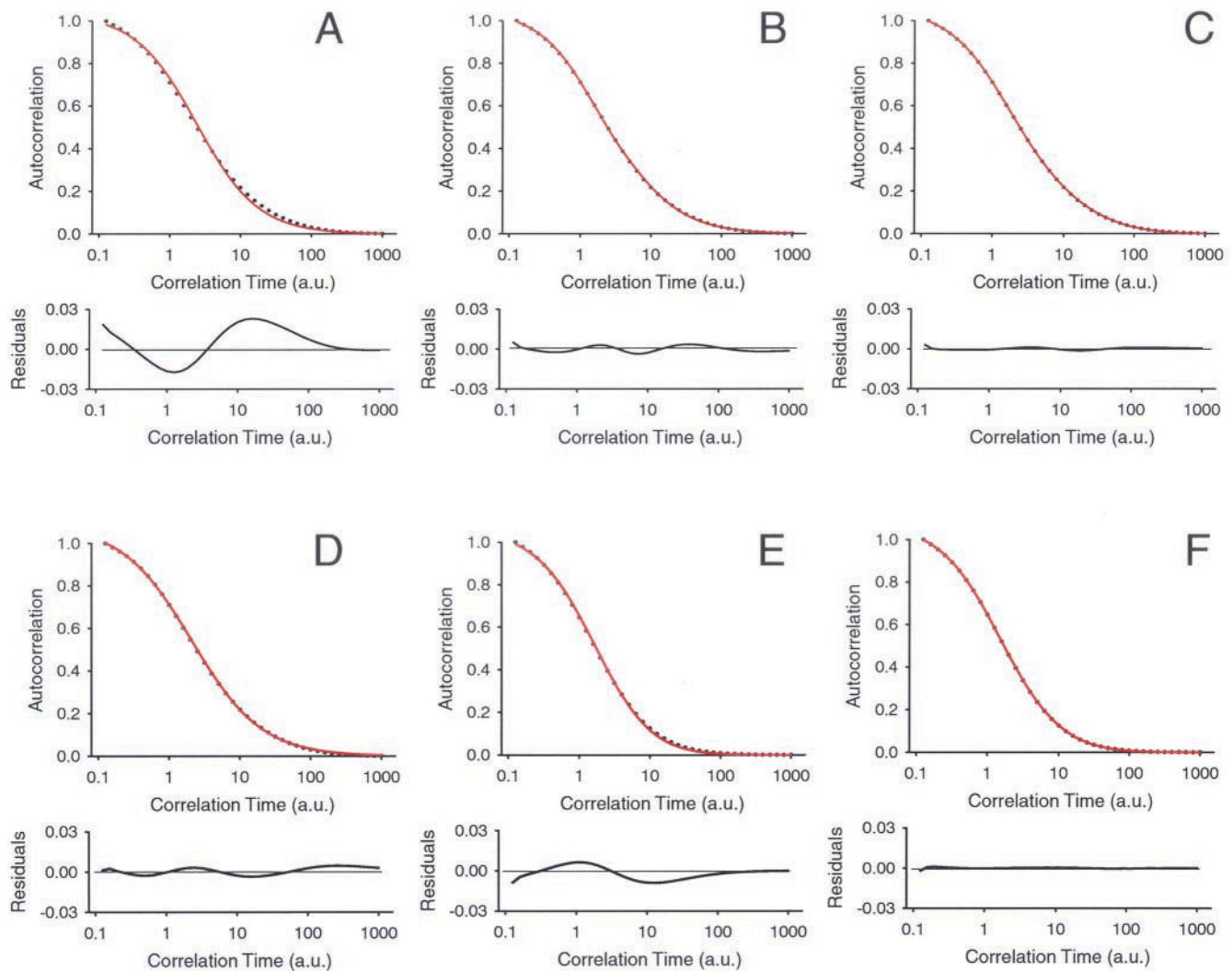
FIGURE 10 Artifacts in the autocorrelation introduced by invalid assumption of a Gaussian observation profile. The diffusion autocorrelation calculated (*black points*) using diffraction theory for a 1.2 NA 40× water objective with overfilled back-aperture ($\beta \ll 1$) and $r_d = 20$ ou, is least-squares fitted using standard FCS analytic functions (*red lines*). Residuals are shown below (*black lines*). (*A*) Fit using Eq. 7 (simple diffusion in a Gaussian observation volume) shows significant residuals. (*B*) Fit with Eq. 12 (diffusion + kinetics) reduced the residuals by introducing an exponential artifact, which could be erroneously interpreted as chemical kinetics. (*C*) Similarly, fit using two diffusing species (Eq. 10) greatly improves residuals but introduces an erroneous interpretation (see text for artifact amplitudes and time scales). (*D*) Using a single-component anomalous diffusion fitting function (Eq. 9) also improves the fit. (*E*) Artifacts result in a larger than normal value for ω. Diffusion autocorrelation for same setup but with smaller detector aperture ($r_d = 4.7$ ou) fit with Eq. 7 with $\omega = 2.5$ (fixed). (*F*) Same as (*E*) with $\omega = 4.3$ (floating for best fit). See text for a detailed account of fitting parameters.

artifacts are more severe ($F_B \sim 0.10$ or larger), allowing the axial ratio to approach artificially large values will improve the fit somewhat, but not eliminate the systematic fitting residuals. Thus, the larger than expected values for ω typically observed are another indication of a non-Gaussian observation volume.

In the case of a poor fit with systematic fitting residuals and a divergent axial ratio, the FCS user is often tempted to use an additional arbitrary exponential fraction or second diffusing component to improve the fit. The altered form of $G(\tau)$ due to a non-Gaussian observation volume can be forced to fit by introducing additional terms in the fitting

function. Additional free parameters, such as an exponential fraction or a second diffusing component, of course improve the fitting residuals, depending on what type of extra factor is added to the fitting function (see Fig. 10).

When the observation volume is strongly non-Gaussian, such as for an overfilled 40×1.2 NA objective with $r_d = 20$ ou, the fitting residuals using Eq. 7 (diffusion-only) are significant and $\omega \rightarrow \infty$ (Fig. 10 *A*). Arbitrarily adding an exponential factor of the form of Eq. 12 greatly improves the fitting residuals (Fig. 10 *B*), yielding erroneous chemical kinetic parameters with notably large amplitude $F_B = 0.31$ and time scale $\tau_B \sim 3.5$ times faster than τ_D, and an increase

in τ_D of 86%. Adjusting the fit by using two diffusing species (Eq. 10, Fig. 10 *C*) also results in an improved fit and misleading fitting parameters: an erroneous second diffusing component ~9 times slower than τ_D with amplitude ~18%. The value of τ_D became 30% slower upon addition of the second species. Because the second diffusing component accommodated the nonstandard form of $G(\tau)$, the axial ratio could approach a smaller value, $\omega \sim 10$. Because these added fitting parameters actually describe the diffusion of a single species in a non-Gaussian volume, the time scale for the artifact will scale with the diffusion time, assuming all other experimental variables are held constant. An anomalous diffusion fit also improved the fitting residuals (Fig. 10 *D*). However, if one chooses to use Eq. 7 and not introduce extra fitting parameters, ω will approach a larger, nonphysical value and improve the fitting residuals when the observation volume is slightly non-Gaussian (Fig. 10, *E* and *F*). While the actual axial ratio of the profile used in the calculation was ~2.5, the fitting parameters yield a best-fit value using Eq. 7 of $\omega \sim 4.3$. These examples illustrate that erroneous fitting parameters will result if the observation volume is non-Gaussian and yet it is fit using any function that assumes a Gaussian form. Errors may occur because of false assumptions about the optical geometry, independent of the properties of the sample; hence, these effects are called artifacts.

As a measure of the severity of the departure of the observation volume from Gaussian and the artifacts that result, the amplitude of an artificial exponential fraction needed to produce a good fit to the diffraction-calculated FCS autocorrelation at various detector apertures and degrees of underfilling is shown in Fig. 11. For all but the least severe artifacts (i.e., $F_B < 0.05$), erroneous chemical and diffusion kinetics will result from fitting the measured autocorrelation using standard fitting functions that assume a Gaussian observation volume.

FCS users should exercise caution in drawing conclusions about any phenomena that might be attributable to such an artifact, particularly when measuring in conditions that are prone to these artifacts: overfilled back-aperture and/or large detector aperture (larger than the r_d that gives maximum η, see below). Conditions that minimize the artifact and provide the most nearly Gaussian observation volume are small detector aperture (typically smaller than the r_d, which gives maximum η), underfilled back-aperture, or two-photon excitation. However, the count rate per molecule is reduced by using a small detector aperture or underfilling the objective. Hence, there is a tradeoff between optimization of signal to noise and achieving a Gaussian observation volume. In all cases a control measurement using a well-characterized, nonaggregating, pure single fluorophore at low intensity *is highly recommended*. Proper calibration under actual measurement conditions is even more important in intracellular FCS, where the systems are inherently more complex, and aberrations and
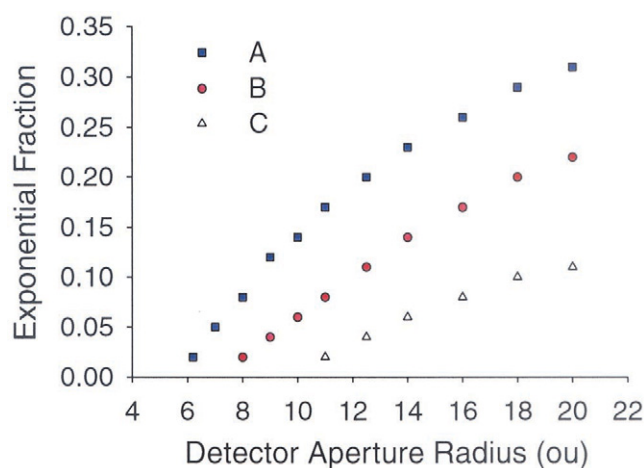

FIGURE 11 The magnitude of the fictitious exponential fractions due to deviations from a Gaussian observation profile. The fraction that produces the best fit between calculated FCS curves and the analytical function is shown versus detector aperture for three underfilling fractions $\beta \ll 1$ (*A*), $\beta = 1.25$ (*B*), and $\beta = 2$ (*C*). These fictitious fractions are artifacts of the optics that are not based on any process or property of the sample, but only on erroneous assumptions about the optical geometry. The artifact is minimized by using an underfilled back-aperture ($\beta > 1$), small detector aperture, or two-photon excitation.

other distortions of the focal volume may be more severe. In the absence of artifacts, the correlation curve can be fit satisfactorily with acceptable signal to noise, using Eq. 7 without an exponential fraction and without diverging values of the axial ratio.

How will other fluorescence fluctuation techniques that use the confocal FCS geometry be able to accommodate a non-Gaussian observation volume? The number of free parameters describing the focal volume is the same in fluorescence intensity distribution analysis (FIDA; Kask et al., 1999) as in FCS. Two parameters describe the observation volume (the axial and radial waist in FCS, and in FIDA the two unconstrained parameters of the three a_k and B(0), the spatial brightness at the origin). Then, for each species there are two additional parameters: concentration and diffusion time (diffusion coefficient) in FCS, and concentration and brightness in FIDA. The two parameters that describe the artifact plus the two that describe the waists make four total in FCS, while the three a_k plus B(0) in FIDA also total four parameters describing the observation volume. Thus, the intrinsic ability of either technique to accommodate non-Gaussian functionality should be the same.

Evidence for a larger illumination profile than predicted: volume, number of molecules, and collected fluorescence

While the diffraction theory does describe the measured autocorrelation better than a Gaussian theory, our description of the observation volume could certainly be improved. Most of the experimental evidence indicates that the actual
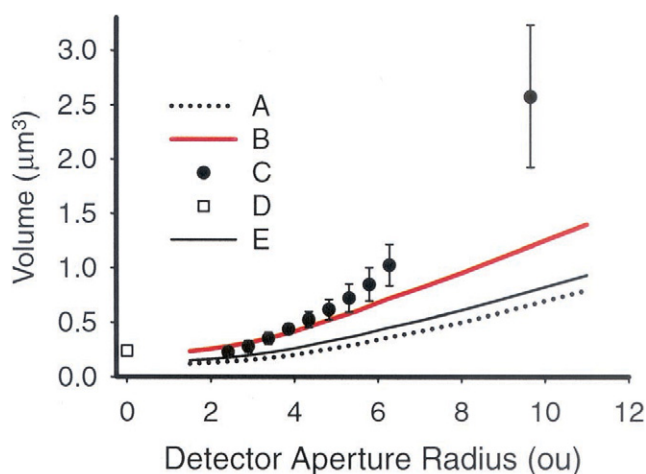
FIGURE 12 Theoretical predictions of the size of the observation volume with (A) $\beta \ll 1$, (B) $\beta = 2$, (E) $\beta = 1.25$. Calculated volumes are compared with measured volumes (C) obtained from FCS measurements of the number of molecules from $G(0)$ and absorbance spectrometer measurements of the concentration, for a 1.2 NA confocal FCS system with detector aperture. Note the steeper slope of the measurements when compared with diffraction theory, particularly at large aperture sizes, evidence that the measured observation volume is larger than predicted. The measured volume with two-photon excitation (D) scaled for the difference in excitation wavelength is shown for comparison.
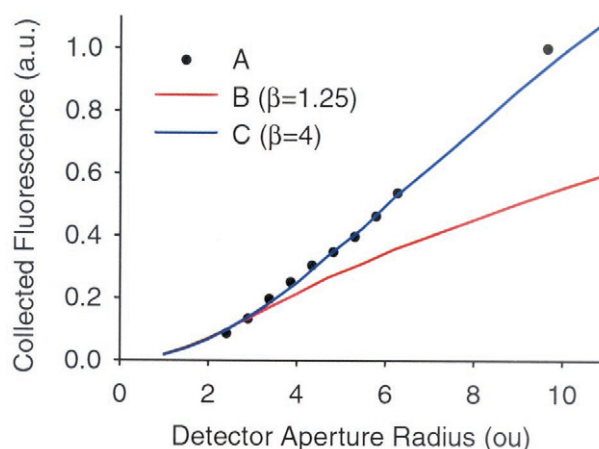

FIGURE 13 Measured 1P-excited fluorescence signal (A) versus detector aperture compared with the diffraction theory with (B) $\beta = 1.25$ and with (C) $\beta = 4$. The theoretical data are multiplied by an arbitrary detection efficiency factor to obtain the best fit to the data. Note that the slope of the measured fluorescence is steeper than theory predicts unless the theoretical value of β is allowed to assume a value much different from the experimental value ($\beta \sim 1$). The deviation becomes most apparent at large detector aperture sizes, where the illumination volume is least constrained by the confocal pinhole (see Discussion).

observation volume is larger than predicted. A direct indicator of this discrepancy is the measured volume of observation, calculated as $V = N/C_0$, (N and C_0 are both measured), which is compared with theoretical predictions (see Fig. 12). Note the measured volume is again larger than the calculated volume, with good agreement only for small apertures. The background from the blank sample was negligible (B $< 0.005 \cdot$ Signal). Although unlikely, collection of background from the fluorescent sample itself cannot be entirely ruled out, which would tend to increase the apparent number of molecules and thus the apparent V value.

Fig. 13 shows the measured collected fluorescence together with the calculated values for $\beta = 1.25$, the actual underfilling fraction, and $\beta = 5$, which corresponds to a larger illumination profile. The calculated curves are scaled by a constant for the best fit with the measured value for a 2.9 ou radius detector aperture. Again, the data are better described by calculations for a larger illumination profile than expected. Therefore, the results consistently indicate that the experimental observation volume is larger than predicted.

Discrepancy between the calculated and actual illumination profile explains the systematic variation in the measured diffusion coefficient of rhodamine green. If the illumination profile and observation profile are larger than predicted, as appears to be the case for large detector apertures, the fit between calculation and measurement will result in a slower diffusion rate than expected and hence a smaller diffusion coefficient for large detector apertures.

Thus, improved modeling of the illumination profile is expected to greatly improve the accuracy of measured diffusion coefficients and other experimental parameters obtained by FCS.

One explanation for an enlarged illumination profile is the presence of optical aberrations (Webb et al., 1990; Wells et al., 1990), which have so far been ignored in this treatment. Previous measurements under similar conditions indicate that the measured illumination profile at the focus of a high-NA objective agrees well with predictions when the objective back-aperture is underfilled ($\beta \sim 5$), but the illumination profile can be significantly larger than predicted ($\sim 30\%$ for $\beta \sim 2.5$) when the objective is more nearly overfilled (Schneider and Webb, 1981). A 30% discrepancy in the waist of the observation volume would lead to nearly a factor of two discrepancy between predicted and measured diffusion coefficient. Due to its high sensitivity and quantitative capabilities, confocal 1P-FCS may be more susceptible to distortions of the illumination profile than 1P confocal microscopy. However, the contribution of aberrations far from the focus would be limited in confocal FCS by the detector aperture, and hence explains the better agreement between measurement and diffraction theory for smaller r_d values. The illumination volume would be enlarged and the contribution of aberrations enhanced if the laser beam is not perfectly collimated upon entry into the objective back-aperture, or if there were lens aberrations affecting the periphery of the focal volume.

However, scattered or reflected fluorescence reaching the detector might be a significant source of background when

using a large detector aperture, and thus would increase the apparent measured number of fluorescent molecules and hence the apparent size of the observation volume. Direct and precise measurement of the illumination profile would answer many of these questions. However, the answers to these questions do not change the result that the actual confocal observation volume is usually non-Gaussian, and this is the basis for most of the conclusions of this paper.

CONCLUSIONS

A non-Gaussian observation volume will result in an altered form for the diffusion autocorrelation function measured by FCS. Unfortunately, this altered functional form can easily mislead FCS users into mistakenly attributing fitting residuals to a chemical kinetic process, a second diffusing species, or some type of nonstandard diffusion. One cause of a non-Gaussian observation volume is the diffraction of light by the apertures within the optical system, particularly the back-aperture of the objective lens. The diffraction fringes are most pronounced when the objective back-aperture is overfilled, and when the confocal detector aperture is opened to the point that it does not confine the observation volume and reduce collection from the fringes. Measurements under typical conditions confirm that the observation volume is non-Gaussian, as evidenced by the peak in count rate per molecule as a function of detector aperture. This peak is also interesting because it signifies the existence of an optimal detector aperture size for maximal signal-to-noise ratio. The best conditions for a Gaussian observation volume are a small detector aperture or underfilled back-aperture, but closing the detector aperture below its optimum value or underfilling the back aperture reduces the signal-to-noise ratio. Thus, the 1P-FCS user is faced with a tradeoff. Two-photon excitation results in a nearly Gaussian observation volume without a detector aperture and provides an advantageous alternative. An improved model of the observation volume that accounts for aberrations might improve agreement between the theory and the 1P measurements. Considering the rapid rate of improvement in computational power, such modeling should be quite feasible in the near future.

REFERENCES

Born, M., and E. Wolf. 1991. Principles of Optics, 6th Ed. Pergamon Press, Oxford. 484.

Brock, R., M. A. Hink, and T. M. Jovin. 1998. Fluorescence correlation microscopy of cells in the presence of autofluorescence. *Biophys. J.* 75:2547–2557.

Brock, R., G. Vamosi, G. Vereb, and T. M. Jovin. 1999. Rapid characterization of green fluorescent protein fusion proteins on the molecular and cellular level by fluorescence correlation microscopy. *Proc. Natl. Acad. Sci. USA.* 96:10123–10128.

Cluzel, P., M. Surette, and S. Leibler. 2000. An ultrasensitive bacterial motor revealed by monitoring signaling proteins in single cells. *Science.* 287:1652–1655.

Cummins, H. Z., and H. L. Swinney. 1970. Light beating spectroscopy. *Prog. in Optics.* 8:135–200.

Eigen, M., and R. Rigler. 1994. Sorting single molecules: application to diagnostics and evolutionary biotechnology. *Proc. Natl. Acad. Sci. USA.* 91:5740–5747.

Elson, E. L., and D. Magde. 1974. Fluorescence correlation spectroscopy. I. Conceptual basis and theory. *Biopolymers.* 13:1–27.

Gu, M., and X. S. Gan. 1996. Effect of the detector size and the fluorescence wavelength on the resolution of three- and two-photon confocal microscopy. *Bioimaging.* 4:129–137.

Gu, M., and C. J. R. Sheppard. 1993. Effects of a finite-sized pinhole on 3D image-formation in confocal two-photon fluorescence microscopy. *J. Mod. Optics.* 40:2009–2024.

Hess, S. T., S. Huang, A. A. Heikal, and W. W. Webb. 2002. Biological and chemical applications of fluorescence correlation spectroscopy: a review. *Biochemistry.* 41:697–705.

Heikal, A. A., S. T. Hess, G. S. Baird, R. Y. Tsien, and W. W. Webb. 2000. Molecular spectroscopy and dynamics of intrinsically fluorescent proteins: coral red (dsRed) and yellow (Citrine). *Proc. Natl. Acad. Sci. USA.* 97:11996–12001.

Hink, M. A., R. A. Griep, J. W. Borst, A. van Hoek, M. H. M. Eppink, A. Schots, and A. J. W. G. Visser. 2000. Structural dynamics of green fluorescent protein alone and fused with a single chain Fv protein. *J. Biol. Chem.* 275:17556–17560.

Kask, P., K. Palo, D. Ullmann, and K. Gall. 1999. Fluorescence-intensity distribution analysis and its application in biomolecular detection technology. *Proc. Natl. Acad. Sci. USA.* 96:13756–13761.

Koppel, D. E. 1974. Statistical accuracy in fluorescence correlation spectroscopy. *Phys. Rev. A.* 10:1938–1945.

Koppel, D. E., D. Axelrod, J. Schlessinger, E. L. Elson, and W. W. Webb. 1976. Dynamics of fluorescence marker concentration as a probe of mobility. *Biophys. J.* 16:1315–1329.

Magde, D., E. Elson, and W. W. Webb. 1972. Thermodynamic fluctuations in a reacting system measurement by fluorescence correlation spectroscopy. *Phys. Rev. Lett.* 29:705–708.

Magde, D., E. L. Elson, and W. W. Webb. 1974. Fluorescence correlation spectroscopy. II. An experimental realization. *Biopolymers.* 13:29–61.

Maiti, S., U. Haupts, and W. W. Webb. 1997. Fluorescence correlation spectroscopy: diagnostics for sparse molecules. *Proc. Natl. Acad. Sci. USA.* 94:11753–11757.

Mertz, J., C. Xu, and W. W. Webb. 1995. Single molecule detection by two-photon-excited fluorescence. *Optics Lett.* 20:2532–2534.

Meseth, U., T. Wohland, R. Rigler, and H. Vogel. 1999. Resolution in fluorescence correlation measurements. *Biophys. J.* 76:1619–1631.

Politz, J. C., E. S. Browne, D. E. Wolf, and T. Pederson. 1998. Intranuclear diffusion and hybridization state of oligonucleotides measured by fluorescence correlation spectroscopy in living cells. *Proc. Natl. Acad. Sci. USA.* 95:6043–6048.

Qian, H., and E. Elson. 1991. Analysis of confocal laser-microscope optics for 3-D fluorescence correlation spectroscopy. *Appl. Optics.* 30:1185–1195.

Richards, B., and E. Wolf. 1959. Electromagnetic diffraction in optical systems. II. Structure of the image field in an aplanatic system. *Proc. R. Soc. Lond. Ser. A.* 253:358–379.

Rigler, R., and E. S. Elson. 2001. Fluorescence correlation spectroscopy: theory and applications. Springer-Verlag, Berlin.

Rigler, R., and J. Widengren. 1990. Ultrasensitive detection of single molecules by fluorescence correlation spectroscopy. *In* BioScience. B. Klinge and C. Owman, editors. Lund University Press, Lund, Sweden. 180–183.

Rigler, R., Ü. Mets, J. Widengren, and P. Kask. 1993. Fluorescence correlation spectroscopy with high count rate and low background-analysis of translational diffusion. *Eur. Biophys. J.* 22:169–175.

Sandison, D. R., D. W. Piston, R. M. Williams, and W. W. Webb. 1995. Quantitative comparison of background rejection, signal-to-noise ratio, and resolution in confocal and full-field laser-scanning microscopies. *Appl. Optics.* 34:3576–3588.

Sandison, D. R., and W. W. Webb. 1994. Background rejection and signal-to-noise optimization in confocal and alternative fluorescence microscopes. *Appl. Optics.* 33:603–615.

Schneider, M. B., and W. W. Webb. 1981. Measurement of sub-micron laser-beam radii. *Appl. Optics.* 20:1382–1388.

Schwille, P., U. Haupts, S. Maiti, and W. W. Webb. 1999a. Molecular dynamics in living cells observed by fluorescence correlation spectroscopy with one- and two-photon excitation. *Biophys. J.* 77:2251–2265.

Schwille, P., J. Korlach, and W. W. Webb. 1999b. Fluorescence correlation spectroscopy with single-molecule sensitivity on cell and model membranes. *Cytometry.* 36:176–182.

Schwille, P., F. Oehlenschlager, and N. G. Walter. 1996. Quantitative hybridization kinetics of DNA probes to RNA in solution followed by diffusional fluorescence correlation analysis. *Biochemistry.* 35: 10182–10193.

Self, S. A. 1983. Focusing of spherical Gaussian beams. *Appl. Optics.* 22:658–661.

Sheppard, C. J. R., C. J. Cogswell, and M. Gu. 1991. Signal strength and noise in confocal microscopy: factors influencing selection of an optimum detector aperture. *Scanning.* 13:233–240.

Sheppard, C. J. R., and H. J. Matthews. 1987. Imaging in high-aperture optical systems. *J. Opt. Soc. Am. A.* 4:1354–1360.

Thompson, N. L. 1991. Fluorescence correlation spectroscopy *In* Topics in Fluorescence Spectroscopy, Vol. 1. Techniques. J. R. Lakowicz, editor. Plenum Press, New York. 337–410.

Visser, N. V., M. A. Hink, A. van Hoek, and A. J. W. G. Visser. 1999. Comparison between fluorescence correlation spectroscopy and time-resolved fluorescence anisotropy as illustrated with a fluorescent dextran conjugate. *J. Fluorescence.* 9:251–255.

Wachsmuth, M., W. Waldeck, and J. Langowski. 2000. Anomalous diffusion of fluorescent probes inside living cell nuclei investigated by spatially-resolved fluorescence correlation spectroscopy. *J. Mol. Biol.* 298:677–689.

Webb, W. W. 1976. Applications of fluorescence correlation spectroscopy. *Q. Rev. Biophys.* 9:49–68.

Webb, W. W., K. S. Wells, D. R. Sandison, and J. H. Strickler. 1990. Criteria for quantitative dynamical confocal fluorescence imaging. *In* Optical Microscopy for Biology. B. Herman and K. Jacobson, editors. Wiley-Liss, New York. 73–108.

Wells, K. S., D. R. Sandison, J. H. Strickler, and W. W. Webb. 1990. Quantitative fluorescence imaging in laser scanning confocal microscopy. *In* Handbook of Biological Confocal Microscopy. J. Pawley, editor. Plenum Press, New York. 27–39.

Wolf, E. 1959. Electromagnetic diffraction in optical systems. I. An integral representation of the image field. *Proc. R. Soc. Lond. Ser. A.* 253:349–357.

Xu, C., and W. W. Webb. 1997. Multiphoton excitation of molecular fluorophores and nonlinear laser microscopy. *In* Topics in Fluorescence Spectroscopy, Vol. 5. J. R. Lakowicz, editor. Plenum Press, New York. 471–540.