Protein Interactions

PROTEIN REVIEWS

Recent Volumes in this Series

VIRAL MEMBRANE PROTEINS: STRUCTURE, FUNCTION, AND DRUG DESIGN
Edited by Wolfgang B. Fischer

THE p53 TUMOR SUPPRESSOR PATHWAY AND CANCER
Edited by General P. Zambetti

PROTEOMICS AND PROTEIN-PROTEIN INTERACTIONS: BIOLOGY, CHEMISTRY, BIOINFORMATICS, AND DRUG DESIGN
Edited by Gabriel Waksman

PROTEIN MISFOLDING, AGGREGATION AND CONFORMATIONAL DISEASES PART A: PROTEIN AGGREGATION AND CONFORMATIONAL DISEASES
Edited by Vladimir N. Uversky and Anthony L. Fink

PROTEIN INTERACTIONS: BIOPHYSICAL APPROACHES FOR THE STUDY OF COMPLEX REVERSIBLE SYSTEMS
Edited by Peter Schuck

PROTEIN MISFOLDING, AGGREGATION AND CONFORMATIONAL DISEASES PART B: MOLECULAR MECHANISMS OF CONFORMATIONAL DISEASES
Edited by Vladimir N. Uversky and Anthony L. Fink

Protein Interactions

Biophysical Approaches for the Study of Complex Reversible Systems

Edited by

Peter Schuck
National Institutes of Health, Bethesda, Maryland

Springer

Peter Schuck
National Institutes of Health
Bethesda, MD 20892
USA

Library of Congress Control Number: 2006928099

ISBN-10: 0-387-35965-6 e-ISBN-10: 0-387-35966-4
ISBN-13: 978-0-387-35965-6 e-ISBN-13: 978-0-387-35966-3

Printed on acid-free paper.

9 8 7 6 5 4 3 2 1

springer.com

Contents

Contributors

Claire A. Adams
Department of Molecular and Cellular Biochemistry
University of Kentucky College of Medicine
Lexington, KY, USA

Peter S. Andersen
Symphogen A/S,
Copenhagen, Denmark

Peter S. Backlund Jr
National Institute of Child Health and Human Development
National Institutes of Health
Bethesda, MD, USA

Hacène Boukari
National Institute of Child Health and Human Development
National Institutes of Health
Bethesda, MD, USA

Christine Ebel
Institut de Biologie Structurale UMR 5075 CEA-CNRS-UJF
Grenoble, France

Dorothy A. Erie
Department of Chemistry
University of North Carolina at Chapel Hill
Chapel Hill, NC, USA

Michael G. Fried
Department of Molecular and Cellular Biochemistry
University of Kentucky College of Medicine
Lexington, KY, USA

Wolfram Gronwald
Institute of Biophysics and Physical Biochemistry
University of Regensburg
Regensburg, Germany

Myun K. Han
Excimus Biotech, Inc.
Baltimore, MD, USA

John J. Harvey
Excimus Biotech, Inc.
Baltimore, MD, USA

Hans Robert Kalbitzer
Institute of Biophysics and Physical Biochemistry
University of Regensburg
Regensburg, Germany

Elizabeth A. Komives
Department of Chemistry and Biochemistry
University of California, San Diego
La Jolla, CA, USA

Jay R. Knutson
National Heart, Lung, and Blood Institute
National Institutes of Health
Bethesda, MD, USA

Werner Kremer
Institute of Biophysics and Physical Biochemistry
University of Regensburg
Regensburg, Germany

John E. Ladbury
Department of Biochemistry & Molecular Biology
University College London
London, UK

Emmanuel Margeat
INSERM U554
CNRS UMR 5048
Montpellier, France

Jacob Piehler
Institute of Biochemistry
Johann Wolfgang Goethe-University
Frankfurt am Main, Germany

Carol V. Robinson
University of Cambridge
Department of Chemistry
Cambridge, UK

Catherine A. Royer
INSERM U554
CNRS UMR 5048
Montpellier, France

Alan M. Sandercock
Department of Chemistry
University of Cambridge
Cambridge, UK

Peter Schuck
National Institute of Biomedical Imaging and Bioengineering
National Institutes of Health
Bethesda, MD, USA

Frank Schumann
Institute of Biophysics and Physical Biochemistry
University of Regensburg
Regensburg, Germany

Michael Spörner
Institute of Biophysics and Physical Biochemistry
University of Regensburg
Regensburg, Germany

Eric J. Sundberg
Boston Biomedical Research Institute
Watertown, MA, USA

Dmitri I. Svergun
European Molecular Biology Laboratory
Hamburg, Germany

Patrice Vachette
Institut de Biochimie et Biophysique Moléculaire et Cellulaire
UMR 8619 CNRS-Université Paris-Sud
Orsay Cedex, France

Hong Wang
National Institute of Environmental Health Sciences
National Institutes of Health
Research Triangle Park, NC, USA

Mark A. Williams
Department of Biochemistry and Molecular Biology
University College London
London, UK

Kristi Wojtuszewski
National Heart, Lung, and Blood Institute
National Institutes of Health
Bethesda, MD, USA

Yong Yang
Department of Chemistry
University of North Carolina at Chapel Hill
Chapel Hill, NC, USA

Alfred L. Yergey
National Institute of Child Health and Human Development
National Institutes of Health
Bethesda, MD, USA

Preface

When I was invited to edit this volume, I wanted to take the opportunity to assemble reviews of different biophysical methodologies for protein interactions at a level sufficiently detailed to understand how complex systems can be studied. There are several excellent introductory texts for biophysical methodologies, many with hands-on descriptions or embedded in general introductions to physical biochemistry. The goal of the present volume was to present state-of-the-art reviews that do not necessarily enable the reader to carry out these techniques, but to gain a deep understanding of the biophysical observables, to stimulate creative thought on how the techniques may be applied to study a particular biological system, and to foster collaboration and multidisciplinary work.

Reversible protein interactions involve noncovalent chemical bonds, producing protein complexes with free energies not far from the order of magnitude of the thermal energy kT. As a consequence, they can be highly dynamic and may be controlled, for example, by protein expression levels and changes in the intracellular or microenvironment. Reversible protein complexes may have sufficient stability to be purified for study, but frequently their short lifetime essentially limits their existence to solutions of mixtures of the binding partners in which they remain populated through dissociation and reassociation processes. To understand the function of such protein complexes, it is important to study their structure and dynamics. Even when these studies take place *in vitro*, they elucidate the principles of the interactions imposed by the protein structures, principles which may be quantitatively modulated but have to be followed *in vivo*.

Maps of protein interaction networks display the interdependence of protein interactions and highlight the importance of interactions of more than two proteins. It is probably safe to assume that we currently know only a small fraction of the protein interactome, in particular, triple or higher-order complexes. Proteins that are able to interact with multiple protein binding partners or other ligands can exhibit higher functionality. This can include, for example, logical switches, with cooperativity steepening the isotherms of binding, resulting in highly sensitive and

discriminate response to a cellular stimulus. In many systems, this involves interplay of multiprotein complexes, binding of small ligands, covalent protein modifications, and conformational changes in proteins.

Techniques to elucidate such linked protein interactions and/or multiprotein complexes, with regard to thermodynamics, kinetics, conformation, and flexibility, and the possible role of spatial confinements to surfaces define the scope of the present volume. The side-by-side presentation of different approaches highlights aspects they have in common and orthogonal viewpoints that may provide opportunities for a synergistic combination. For example, an important recurring theme is the role of protein solvation, which is addressed in many chapters from different perspectives. To illustrate the use of the techniques, some applications are described, which, at the same time, are also aimed at providing a kaleidoscopic view of different biological systems and principles of protein interactions.

The list of biophysical methods reviewed in the present volume is far from complete. The selection of topics should not be understood in the sense of merit or importance of the different methods, but rather is a reflection of practical limitations in the scope and assembly of this work.

I want to thank the authors for their contributions. I also want to thank the series editor Dr. Zou Atassi, the publisher, and the National Institutes of Health for making this work possible. Finally, I want to thank my beloved wife Teresa for her patience and support.

Bethesda, 2005

1

The Characterization of Biomolecular Interactions Using Fluorescence Fluctuation Techniques

Emmanuel Margeat, Hacène Boukari, and Catherine A. Royer*

1.1. INTRODUCTION

Fluorescence correlation spectroscopy (FCS) has become one of the most popular methods available for investigating the physical properties of biomolecular complexes ever since it was first proposed by the groups of Elliot Elson and Watt Webb in the 1970s (Aragon and Pecora, 1974; Magde *et al*., 1974; Webb, 1976; Icenogle and Elson, 1983; Elson, 2001). Over the years, a number of thorough reviews have appeared of the now extensive literature (Madge, 1976; Webb, 1976; Berland *et al*., 1995; Rigler, 1995; Schwille, 2001; Hess *et al*., 2002; Thompson *et al*., 2002; Haustein and Schwille, 2003; Muller *et al*., 2003; Elson, 2004; Haustein and Schwille, 2004). While the early work in the field was carried out on dyes and purified biological molecules, it became quickly apparent that FCS was well suited for cellular applications as well (Berland *et al*., 1995; Schwille, 2001; Hess *et al*., 2002; Krivensky and Bonnet, 2002; Bacia and Schwille, 2003; Vukojevic *et al*., 2005). FCS has been used to study lipid diffusion and protein associations in model and biological membranes, nucleic acid hybridization and protein–nucleic acid interactions, protein–protein interactions, both homologous and heterologous, in solution and in cells, and protein and DNA associations with

E. MARGEAT • INSERM, U554, Montpellier, 34090 Cedex, France; CNRS-UM-UM2, UMR5048. H. BOUKARI • National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD, USA. C. A. ROYER • INSERM, U554, Montpellier, 34090 Cedex, France; CNRS-UM1-UM2, UMR5048 and *Corresponding author: *e*-mail: royer@cbs.cnrs.fr.

small ligands. There are literally hundreds of articles in the PubMed database on the subject of the use of FCS to study biomolecular interactions. Therefore, it is outside the scope of the present work to review this abundant literature in detail. Rather, this review touches on some of the more practical aspects of the use of FCS in binding studies. The advantages vis-à-vis other techniques are discussed, and some examples of the applications of FCS using simple diffusion time measurements, photon statistics, and cross-correlation measurements are presented. Finally, some of the more problematic artifacts are described, along with approaches designed to minimize their contributions.

We restrict the discussion here to experimental setups in which the fluorescence signal from diffusing molecules is detected in a very small, open volume at concentrations that are sufficiently low such that significant fluctuations in the fluorescence intensity are detected as molecules enter and leave the volume, i.e., at or near the single molecule level. In general, an FCS setup (see Figure 1.1) is based on laser excitation into an inverted microscope equipped with an objective featuring a high numerical aperture. Detection is accomplished through dichroic mirrors and appropriate filters by photocounting detectors such as avalanche photodiodes or photomultiplier tubes. The time sequence of the detected intensity is either directly processed by a fast digital autocorrelator (hardware) or stored as a raw photon versus time stream and processed using autocorrelation software.

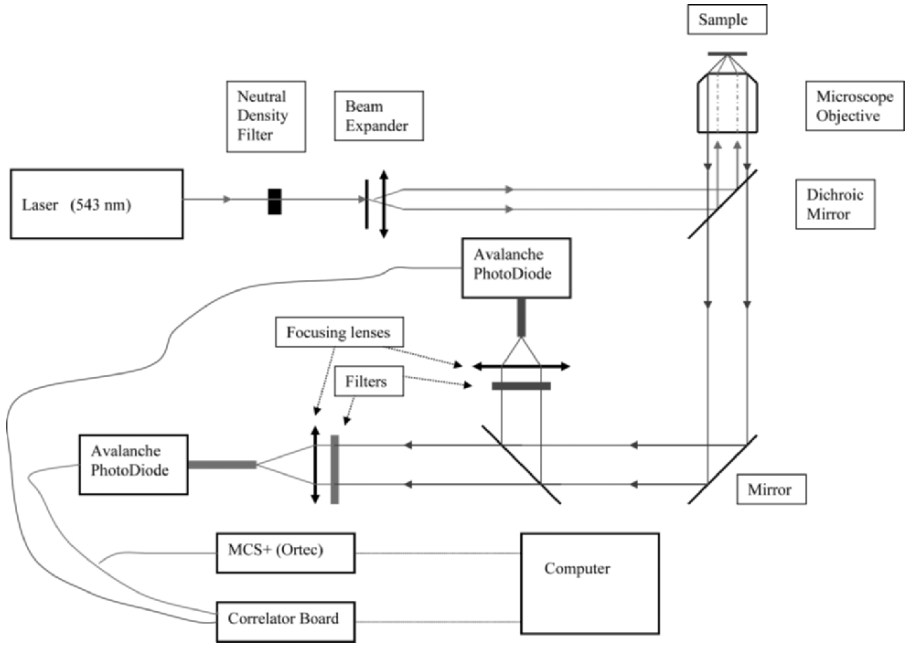

Figure 1.1. Schematic diagram of a one-photon FCS system.

Examples of various experimental setups can be found in the reviews cited earlier. The small open volume (typically <1 fL) can be achieved by either using visible laser excitation coupled with a confocal pinhole (Eigen and Rigler, 1994) or small diameter optical fibers (10–50 μm) to limit the detection volume, or by two-photon excitation with a femtosecond pulsed IR laser beam through a high numerical aperture objective (Denk *et al*., 1990), which focuses the beam to a diffraction-limited spot with a diameter equal to $\sim \lambda/2$. The point spread function is typically considered to exhibit a 3-D Gaussian profile. In two-photon excitation (TPE), molecules simultaneously absorb two photons whose combined energy allows the transition to an excited electronic state. While in one-photon excitation, the emission is at longer wavelengths than the excitation, in TPE, the emission is of higher energy than the excitation. The probability of TPE depends on the square of the incident photon flux. For a discussion of the two-photon excitation theory see the review by So *et al*. (2000) and references therein.

1.2. ADVANTAGES OF FLUCTUATION SPECTROSCOPY

FCS and related techniques benefit from the fundamental advantage that fluorescence in general presents over a number of other approaches to the quantitative study of biomolecular interactions, namely its high sensitivity. This high sensitivity allows for the fast detection of emission from as little as a single molecule. Fast detection renders accessible the analysis of events on fast timescales. The ultimate lower limit of the timescale for fluorescence detection is set by the rate at which emission occurs, typically on the nanosecond timescale. Another important advantage of fluorescence, in principle, is that it can be measured in solution, in gels, or in live cells. Since the observable is fluorescence, the background physical properties of the medium are much less important than in other techniques, such as dynamic light scattering (DLS), for example. Conceptually, DLS and FCS use a similar approach, namely time-correlating a fluctuating signal, which is further analyzed to extract the size of the macromolecules. Both techniques are relatively noninvasive as they use the interaction of an optical beam with the bio-macromolecules of interest. The underlying physical mechanism for these fluctuations is, however, different for each technique. As noted previously, FCS fluctuations are attributed to changes in intensity induced by the fluorescent bio-macromolecules moving in and out of an excitation volume (we ignore the internal photophysical fluctuations). In contrast, the scattered intensity fluctuations in DLS result from the constructive and destructive interference of the electric fields emanating coherently from the bio-macromolecules present in a small probed volume. Each technique has its own advantages and limitations. As described earlier, FCS requires nanomolar concentration of fluorescent bio-macromolecules, which could be an important factor in bioassays of scarce materials, whereas micromolar or higher concentrations are generally needed in DLS experiments. Because of fluorescence, the bio-macromolecules of interest can be specifically distinguished from a host of other particles, hence allowing the study of

their dynamical behavior in the presence of the other particles. This is the main reason that FCS has become attractive to cellular studies, where the cells are notoriously known for their crowded environment. In DLS all the particles present in the sample are potential scatterers and contribute to the scattered and detected signal, making it challenging to analyze the resulting correlation function and to extract the characteristics of the macromolecules of interest. An advantage of DLS is the possibility of studying dynamical processes at various length-scales by performing angle-dependence measurements; the angle, commonly defined as the scattering angle, subtends the direction of the incident beam and that of detection. In FCS, there is practically one length-scale, limited by the diffraction limit of the focused beam. Finally, the macromolecule of interest does not need to be labeled in DLS experiments, whereas FCS requires a genetic, enzymatic, or chemical labeling of the studied macromolecule.

Consequently, fluorescence allows for the specific observation of the fluorescent species in relatively heterogeneous media, which is advantageous even in *in vitro* experiments. It allows one to monitor associations involving the fluorescent species in the presence of unlabeled competitor molecules, for example. Multicolor experiments can give access to the separate observation of the interactions of different proteins, with each other or competitors. Moreover, this specificity of observation allows detection in a background such as gels and cells, in which many other molecules with diverse physical properties are present.

When measuring molecular interactions by the observation of some fluorescence signal (in the present case, intensity fluctuations) there is no need to separate bound from free species because their respective signals can be differentiated in the overall fluorescence parameter. This means that fluorescence fluctuation measurements can be made under conditions of true equilibrium, with target molecules present at concentrations usually well below the dissociation constant for the complex. Measurements in solution also allow for the relatively simple modulation of solution conditions, such as temperature, concentration of interacting species, pH, ionic strength, and ligand concentration. Thus the role of these parameters in modulating the interaction can be assessed in a straightforward manner.

Fluctuation spectroscopy presents another practical advantage over a number of other techniques used in the study of biomolecular complexation, its relatively small requirement for sample. For example, given the fact that the excitation volumes involved are $\sim$1 fL or less, the volume of sample required for FCS measurements is even smaller than that required for classical fluorescence studies, in the 10-μL range for practical reasons. Such a small observation volume can be achieved using a sandwich of two coverslips separated by a silicone isolator (Molecular Probes, Eugene, OR) for example. The single molecule sensitivity means that the fluorescently labeled molecule is present at nanomolar concentrations or lower in these small volumes, which for a protein of 50 kDa corresponds to less than 0.5 ng of material. Such low levels of materials allow access to dissociation constants for high affinity interactions not accessible to analytical ultracentrifugation and calorimetric approaches. They also render possible experiments that otherwise cannot be carried out for lack of enough material.

Fluctuation spectroscopy can also complement fluorescence anisotropy-based studies of protein–protein, protein–ligand, or protein–nucleic acid interactions. The size range of accessible biomolecular complexes in measurements of the rotational correlation time based on fluorescence anisotropy is limited to correlation times about tenfold longer than the fluorescence lifetime ($\tau_c \sim 40$–80 ns or MW sphere $\sim$100–200 kDa). This is because significant rotation must occur during the excited state lifetime, to reliably determine the rotational correlation. However, there is no theoretical upper limit in fluctuation spectroscopy, other than that associated with diffusion being so slow that photobleaching becomes a problem. While the translational diffusion coefficient is less sensitive to molecular size than is the rotational diffusion coefficient, we shall see subsequently that the resolution levels are relatively high. Moreover, while the interpretation of both time-resolved and steady-state anisotropy is complicated by the need to distinguish local probe mobility from global macromolecular tumbling, the local mobility of the probe is not an issue in the measurement of translational diffusion or molecular brightness by fluctuation spectroscopy. Finally, since fluctuation spectroscopy is also a particle-counting technique, even when molecular weight changes are small, one can rely on changes in the molecular brightness of complexes to monitor interactions.

1.3. FLUORESCENCE CORRELATION SPECTROSCOPY, AND MOLECULAR DIFFUSION

The general theory of fluorescence fluctuation spectroscopy has been presented at length in the reviews cited earlier. Therefore, in the present review only a succinct introduction to FCS theory is given and we limit our discussion to number fluctuations, since these are most useful in the study of biomolecular interactions. We note that there exists a wide range of applications of fluctuation spectroscopy based on the measurement of fluctuations due to changes in conformation (through FRET or quenching). These measurements allow for the characterization of the amplitude and timescale of these molecular motions and have been used in the study of protein folding, enzyme kinetics, and other conformational fluctuations. These molecular phenomena are not the subject of the present review, and are not treated here. We refer the reader to the following articles for further information on this subject (Kettling *et al*., 1998; Grunwell *et al*., 2001; Chattopadhyay *et al*., 2002; Cotlet *et al*., 2004; Joo *et al*., 2004; Li *et al*., 2004; Chattopadhyay *et al*., 2005; Karymov *et al*., 2005; Lee *et al*., 2005; Li and Yeung, 2005; Sato *et al*., 2005; Schuler, 2005; Slaughter *et al*., 2005; Wilson *et al*., 2005).

For those interested in a more detailed introduction to correlation spectroscopy, the recent review by Haustein and Schwille (Haustein and Schwille, 2004) provides a clear presentation of FCS theory. Briefly, if we consider typical measurements of fluorescence intensity of a sample of fluorescent molecules excited in a small illuminated volume, one can define the fluorescence fluctuations as $\delta F(t) = F(t) - \langle F \rangle$, where $F(t)$ is the fluorescence intensity measured at time,

t, and $<F>$ denotes the time-averaged fluorescence intensity. These fluctuations are then time-correlated to generate an autocorrelation function $G(\tau)$, defined as

$$G(\tau) = \frac{\langle \delta F(t)\ \delta F(t+\tau) \rangle}{\langle F \rangle^2} \tag{1.1}$$

with τ being the lag time. The intensity fluctuations are assumed to be directly related to fluctuations in the concentration of the fluorescent molecules in the illuminated volume and can be expressed as

$$\delta F(t) = A \int W(\vec{r})\ \delta c(\vec{r}, t) \mathrm{d}\vec{r}, \tag{1.2}$$

where $W(\vec{r})$ denotes the profile of the excitation volume (usually the laser-beam profile), δc the concentration fluctuation around the average concentration, and A, a constant. The concentration fluctuations are induced by a number of mechanisms, the most studied one being the diffusion of fluorescent molecules in and out of the small excitation volume. For an ideal case of monodisperse, uniformly bright, and freely diffusing fluorescent molecules, a closed-form expression for Eq. (1.1) was derived

$$G(\tau) = \frac{1}{N}\left(1 + \frac{\tau}{\tau_\mathrm{d}}\right)^{-1}\left(1 + \frac{r_\mathrm{o}^2 \tau}{z_\mathrm{o}^2 \tau_\mathrm{d}}\right)^{-1/2}, \tag{1.3}$$

which is commonly used to analyze measured autocorrelation functions, more precisely extract two parameters: the diffusion time, τ_d, and the average number of molecules, N. Here it is assumed that the fluorescent molecules are excited by a 3D Gaussian beam such that,

$$W(r, z) = \mathrm{B}\mathrm{e}^{(-2\mathrm{r}^2/\mathrm{r}_\mathrm{o}^2)}\mathrm{e}^{(-2\mathrm{z}^2/\mathrm{z}_\mathrm{o}^2)} \tag{1.4}$$

characterized by r_o and z_o, respectively, the $1/e^2$ Gaussian intensity beam waists in the radial and axial dimensions as defined by the direction of the laser beam. The extrapolated values of the autocorrelation function in Eq. (1.1) at $\tau = 0$ can be rewritten as

$$G(0) = \frac{1}{\langle C \rangle V_\mathrm{eff}}, \tag{1.5}$$

where $V_\mathbf{eff}$ is an effective observation volume (Nagy and Schwabe, 2004b). Generally, this volume is expressed in terms of that derived from the intensity profile (the point-spread function) $V_\mathbf{psf}$ such that

$$G(0) = \frac{\gamma}{\langle C \rangle V_\mathrm{psf}} \tag{1.6}$$

with

$$V_\mathrm{psf} = \frac{\int W(\vec{r})\ \mathrm{d}\vec{r}}{W(0)}$$

being a mathematical normalization factor for the chosen excitation profile. For example, in the 3D Gaussian profile we have

$$V_{\mathrm{psf}} = \left(\frac{\pi}{2}\right)^{3/2} (r_{\mathrm{o}}^2 z_{\mathrm{o}}). \tag{1.7}$$

In Eq. (1.6) the γ factor provides a measure of the uniformity of the fluorescence intensity observed for molecules at different positions within the volume and the abruptness of the boundaries of the latter (Nagy and Schwabe, 2004a). The γ factor is typically less than 1; that is, the effective observation volume is larger than the volume of the spread function. The γ factor is only equal to 1 for a true physical volume with well-defined boundaries.

In principle, one can determine the effective volume, V_{eff}, from the intercept ($\tau \rightarrow 0$) of measured correlation functions from standard fluorophores with known concentration. In practice, however, possible experimental artifacts from both the studied sample and the instrument, which are discussed later, need to be taken into account for an absolute calibration. Thereafter, for unknown samples, in addition to the diffusion time, the concentration, C, of the fluorescent species in solution can also be determined.

The second relevant parameter in Eq. (1.3) is the diffusion time, τ_{d}, which is related to the translational diffusion coefficient as follows for one- and two-photon excitations, respectively,

$$\tau_{\mathrm{d}} = \frac{r_{\mathrm{o}}^2}{4D} \quad \text{and} \quad \tau_{\mathrm{d}} = \frac{r_{\mathrm{o}}^2}{8D} \tag{1.8}$$

the factor 8 in the denominator of the second expression arising because of the quadratic dependence of fluorescence intensity on excitation intensity in two-photon excitation. Furthermore, in dilute solutions, one can apply the Stokes–Einstein relation of the diffusion coefficient, D,

$$D = \frac{k_{\mathrm{b}} T}{6\pi\eta r_{\mathrm{h}}} \tag{1.9}$$

to determine the hydrodynamic radius, r_{h}. In Eq. (1.9) k_b denotes the Boltzman constant ($1.38 \times 10^{-23}\,\mathrm{kg\,m^2/s^2/K}$), T is the temperature in Kelvin, and η is the the solvent viscosity. The viscosity of water at 20°C is ~1 cp (10^{-3} Pa s or 10^{-3} kg / (m × s). The radius of spherical molecules is the hydrodynamic radius, r_{h}, which can be related to the molecular weight, M (g/mol), Avogadro's number, N (molecules/mol), and the hydrated volume, V_{h}, of the protein ($\sim 1.03\,\mathrm{cm^3/g}$).

$$r_{\mathrm{h}} = \left(\frac{3MV_{\mathrm{h}}}{4\pi N}\right)^{1/3}. \tag{1.10}$$

For a spherical protein of 10 kDa, the hydrodynamic radius is 1.5 nm, while for a protein of 100 kDa, r_{h} is 3.4 nm. Substituting the value of r_{h} into Eq. (1.9) yields

diffusion coefficients of $143 \times 10^{-12}\,m^2/s$ or $143\,\mu m^2/s$ for the 10 kDa spherical protein and $63\,\mu m^2/s$ for the 100 kDa spherical protein. Thus, the smaller the fluorescent molecule, the faster will be its diffusion, and the autocorrelation function will decay to zero more quickly (e.g., Figure 1.2). However, the function is not a strong one owing to the cube root dependence of the radius of the protein on the molecular weight. Hence a factor difference of 10 in size yields only a little more than a factor difference of 2 in the diffusion coefficient.

An intuitive picture of the autocorrelation function can be imagined as follows. If the molecule is large, then the fluorescence signal remains high (similar to that at time, *t*) for a longer period than if the molecule is small, since the larger molecule takes a longer time to leave the volume once it has entered. Thus, the amplitude of the autocorrelation function remains high for a longer period before decaying to zero.

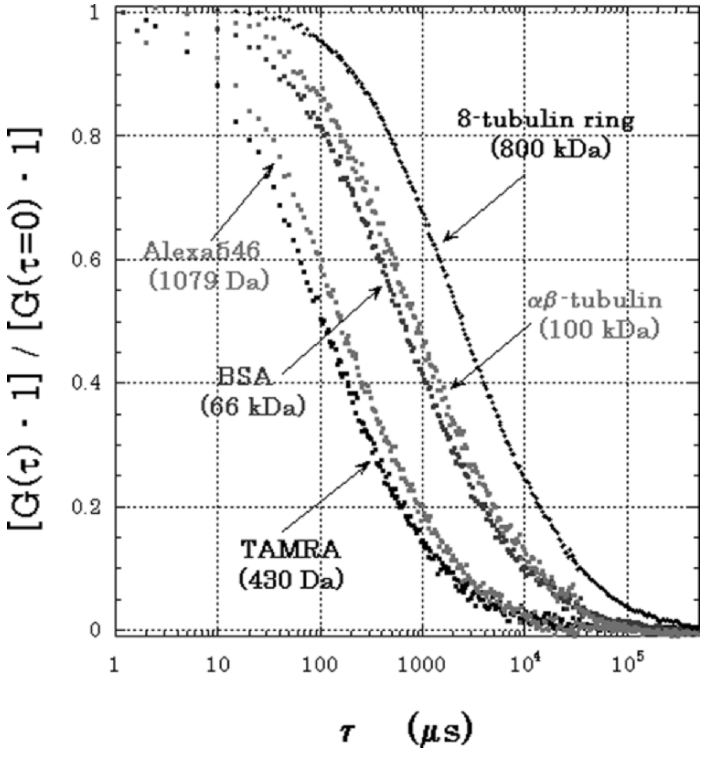

Figure 1.2. One-photon autocorrelation profiles of TAMRA (0.43 kDa), Alexa 546 (1.079 kDa), TAMRA - labeled BSA (66 kDa), the TAMRA - labeled αβ tubulin dimer (100 kDa), and the TAMRA - labeled 8 tubulin ring (800 kDa). It can be seen that a difference of much less than a factor of 2 (66 and 100 kDa) can be resolved if care is taken in the measurements.

Moreover, since the amplitude of the correlation is inversely proportional to the number of molecules, the higher the concentration used in the experiment, the lower the maximum value [or $G(0)$] of the autocorrelation function. This relationship can be intuitively understood by considering that the entrance or exit of a fluorescent particle into or from the detection volume when no other fluorescent molecules are present gives rise to a large relative fluctuation, whereas, if the concentration is high, such that many molecules are present on average in the volume, then the fluctuation due to the entrance or exit of a single molecule is quite small.

Autocorrelation functions for a fluorescent species interacting with itself or another nonfluorescent species can be analyzed in terms of the diffusion times of the bound and free species, their molecular brightness values, and their respective fractional populations.

$$G(\tau) = \frac{1}{N}\left(\left(Y_{\mathrm{f}}\left(1+\frac{\tau}{\tau_{\mathrm{df}}}\right)^{-1}\left(1+\frac{r_{\mathrm{o}}^2\tau}{z_{\mathrm{o}}^2\tau_{\mathrm{df}}}\right)^{-1/2}\right)+\left(Y_{\mathrm{b}}\left(1+\frac{\tau}{\tau_{\mathrm{db}}}\right)^{-1}\left(1+\frac{r_{\mathrm{o}}^2\tau}{z_{\mathrm{o}}^2\tau_{\mathrm{db}}}\right)^{-1/2}\right)\right), \tag{1.11}$$

where

$$Y_{\mathrm{f}} = \frac{\varepsilon_{\mathrm{f}}^2 X_{\mathrm{f}}}{(\varepsilon_{\mathrm{f}} X_{\mathrm{f}} + \varepsilon_{\mathrm{b}} X_{\mathrm{b}})^2} \text{ and } \quad Y_{\mathrm{b}} = \frac{\varepsilon_{\mathrm{b}}^2 X_{\mathrm{b}}}{(\varepsilon_{\mathrm{f}} X_{\mathrm{f}} + \varepsilon_{\mathrm{b}} X_{\mathrm{b}})^2} \tag{1.12}$$

with ϵ_{f} and ϵ_{b}, τ_{df} and τ_{db}, X_{df} and X_{db} corresponding to the molecular brightness diffusion time and the fractional population of the free and bound species, respectively. N here represents an effective number of total molecules. If the brightness of the free and bound species is the same, then Y_{f} and Y_{b} correspond directly to the fractional populations of the two species. In carrying out such experiments it is important to use a global analysis of the entire family of autocorrelation curves for the multiple concentrations tested in order to recover the diffusion times, molecular brightness values, and populations with a reasonable degree of certainty.

1.4. CROSS-CORRELATION AND HETEROLOGOUS ASSOCIATIONS

While the interpretation of the autocorrelation function can be hindered by anomalous or hindered diffusion, or small changes in molecular weight, two-color cross-correlation experiments provide a much clearer indication of interaction between biomolecules. The previous discussion of the autocorrelation function pertains to a single-channel correlation instrument. However, if one can arrange to label two partners in a biomolecular interaction with two different fluorophores that emit at different wavelengths, to excite simultaneously in time and in space the two fluorophores, and finally to simultaneously detect their emission on two separate channels, i and j, then it is possible to cross-correlate the traces of intensity versus time of from the two channels. These fluorophores should be chosen so as to avoid any energy transfer since FRET is anti-correlated, and also any cross-talk

(also referred to as bleed-through) of the emission between the two channels. The amplitude of this cross-correlation signal is directly related to the degree of interaction between the two labeled molecules (Figure 1.3). If the two dyes are in the same complex, then as they diffuse in and out of the observation (or excitation volume) the fluctuations in their intensity will be correlated, whereas if the two biomolecules do not interact then the intensity fluctuations of their two dyes due to their diffusion in and out of the effective volume will have no relationship to each other in time.

We consider here, as earlier, only intensity fluctuations due to number fluctuations. Under ideal conditions in which there is no cross-talk and the intensity measured in channel i only emanates from species 1 and 12, while that in channel j only arises from species 2 and 12, the cross-correlation function then becomes

$$G_{ij}(\tau) = \frac{\langle F_i(t)\, F_j(t+\tau)\rangle}{\langle F_i(t)\rangle\, \langle F_j(t)\rangle}, \tag{1.13}$$

$$G_{ij}(\tau) = \frac{(\langle C_{12}\rangle\, M_{12}(\tau))}{V_{\text{eff}}(\langle C_1\rangle + \langle C_{12}\rangle)\,(\langle C_2\rangle + \langle C_{12}\rangle)}, \tag{1.14}$$

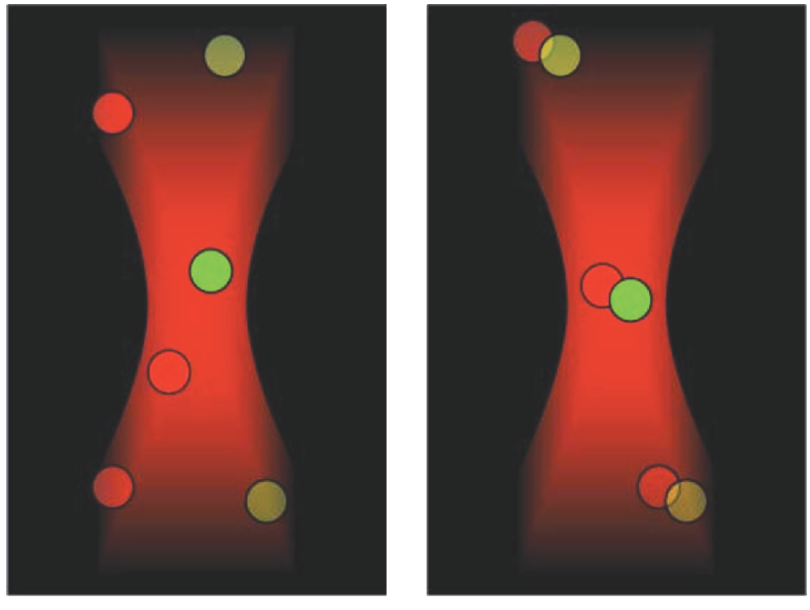

Figure 1.3. Schematic representation of cross-correlation studies of a heterologous protein–protein interaction. One protein here is labeled with a green dye and the other with a red dye. The colors of the dyes are chosen to avoid cross-talk in the detection channels. The fluctuations in fluorescence intensity due to self-diffusion in the detection volume detected on two separate channels are only correlated with each other in time if the two proteins are in interaction (*right*).

where C_1, C_2, and C_{12} are the concentrations of the free and interacting species, respectively, and M_{12} is the term describing the diffusion of the complex.

$$M_{12} = \left(1 + \frac{\tau}{\tau_{d12}}\right)^{-1} \left(1 + \frac{r_o^2 \tau}{z_o^2 \tau_{d12}}\right)^{-1/2}. \tag{1.15}$$

The autocorrelation functions from each of the channels can be expressed as

$$G_{ii}(\tau) = \frac{(\langle C_{11}\rangle M_{11}(\tau) + \langle C_{12}\rangle M_{12}(\tau))}{V_{eff}(\langle C_{11}\rangle + \langle C_{12}\rangle)},$$
$$G_{jj}(\tau) = \frac{(\langle C_{22}\rangle M_{22}(\tau) + \langle C_{12}\rangle M_{12}(\tau))}{V_{eff}(\langle C_{22}\rangle + \langle C_{12}\rangle)}. \tag{1.16}$$

If the two fluorescent species are noninteracting and hence do not diffuse together, then the fluctuations in intensity due to their number fluctuations will be entirely uncorrelated. In this case, the M_{12} term is null and the cross-correlation function becomes zero at all times. If on the other hand, the two fluorescent species are in complex with each other, then their fluctuations will be 100% correlated, and the amplitude of the cross-correlation function will reach that of the lower of the two autocorrelation functions.

Simultaneous excitation of two fluorophores can be achieved either using two co-axially aligned laser beams of different wavelengths (Kettling *et al*., 1998) or alternately through two-photon excitation using a pulsed IR laser at a single wavelength (Heinze *et al*., 2000). The broad two-photon cross-sections of many fluorophores are due to the different selection rules for the two-photon transition and allow for simultaneous excitation of fluorophores that exhibit large wavelength differences in their one-photon absorption spectra. Two-photon excitation for two-color cross-correlation is easier to align since there is only one excitation source and no emission pinholes. Thus the observation volume is exactly the same for the two detection channels. In one-photon excitation, the two excitation lasers must be exactly aligned so that the excitation volume is equivalent. Moreover, if each of the detectors has its own emission pinhole, then these must be perfectly aligned such that the focal volume defined by the two pinholes is exactly the same. Alternately, only one pinhole can be used before splitting the emission between the two channels.

The equations above pertain to ideal conditions in which the singly labeled species are detected only in their respective channels, i.e, there is no cross-talk, and only the doubly labeled species are detected in both channels. Otherwise, one must take into account the contributions of all of the species in each of the channels. Considering three species 1, 2, and 12 corresponding to the free species labeled with a short wavelength emitting dye, the free species labeled with a long wavelength emitting dye, and the complex between the two, then in channel i, the fluorescence fluctuations will include contributions from species 1 of concentration C_1 with ε_{1i}, its brightness in channel i, from species 2 of concentration C_2 with ε_{2i}, its brightness in channel i, and species 12 of concentration C_{12} with ε_{12i}, its brightness in channel i.

Similar contributions will hold in channel j. Thus, the amplitude in channel i (or j) at time zero can be expressed as follows:

$$G_i(0) = \frac{(\varepsilon_{i1}^2 C_1 + \varepsilon_{i2}^2 C_2 + \varepsilon_{i12}^2 C_{12})}{V_{\text{eff}}(\varepsilon_{i1} C_1 + \varepsilon_{i2} C_2 + \varepsilon_{i12} C_{12})^2} \tag{1.17}$$

while the amplitude of the cross-correlation function at time zero will be

$$G_{ij}(0) = \frac{(\varepsilon_{i1}\varepsilon_{j1} C_1 + \varepsilon_{i1}\varepsilon_{j2} C_2 + \varepsilon_{i12}\varepsilon_{j12} C_{12})}{V_{\text{eff}}(\varepsilon_{i1} C_1 + \varepsilon_{i2} C_2 + \varepsilon_{i12} C_{12})(\varepsilon_{j1} C_1 + \varepsilon_{j2} C_2 + \varepsilon_{j12} C_{12})}. \tag{1.18}$$

Indeed, while the interpretation of the time-dependent part of the cross-correlation poses the same difficulties as that for autocorrelation, the analysis of the amplitude of the cross-correlation function at time zero allows for a quantitative determination of the degree of interaction, even in the presence of cross-talk.

Two-color two-photon fluorescence cross-correlation spectroscopy (TCTPFCCS) was first applied to the study of biomolecular interactions *in vitro* (Kettling *et al*., 1998; Heinze *et al*., 2000; Rippe, 2000; Rarbach *et al*., 2001; Heinze *et al*., 2002; Kohl *et al*., 2002; Berland, 2004; Heinze *et al*., 2004; Jahnz and Schwille, 2005). In addition to having contributed significantly to the use of this technique for *in vitro* applications, Schwille's group has recently pioneered its use in live cells, as well. Since they first measured cholera toxin subunit interactions after endocytosis by TCTPFCCS (Bacia *et al*., 2002), several applications either based on fluorescent protein fusions or microinjections of labeled proteins have appeared recently in which quantitative measurements of protein–protein interactions have been made in live cells using this approach (Kim *et al*., 2004; Kim *et al*., 2005; Kohl *et al*., 2005; Larson *et al*., 2005). The two-photon cross-sections allow for simultaneous excitation of any pair of fluorescent proteins (FPs) at wavelengths between 800 and 1,000 nm (Blab *et al*., 2005), which can be achieved using tunable femtosecond IR lasers. Applications in live cells based on simultaneous two-color excitation with visible lasers have been published recently, as well (Vamosi *et al*., 2004; Baudendistel *et al*., 2005; Sato *et al*., 2005; Thews *et al*., 2005). As noted earlier, it is best to work under conditions in which there is no cross-talk between channels, and indeed it is important to avoid FRET between the two dyes, since FRET is anti-correlated. The most appropriate fluorescent protein pairs that would fit these requirements are eCFP or monomeric Cerulean with a red fluorescent protein. Recently monomeric forms of this latter have become available (Campbell *et al*., 2002), allowing for the study of AP1 subunit interactions in live cells in the absence of any cross-talk (Baudendistel *et al*., 2005). Finally, one should note that cross-correlation with a single visible laser is also possible using fluorophores such as quantum dots, which despite similar excitation profiles exhibit distinct emission maxima (Hwang and Wohland, 2005; Swift *et al*., 2006).

We note that while two-color cross-correlation measurements are best used for studying heterologous associations, since the partners can be separately

labeled with different dyes, it is also possible to use the technique for homologous oligomerization studies. While the dynamic range of the value of $G(0)$ is limited to 50% of the total change that could be observed for a heterologous interaction, mixing equal amounts of two preparations of the same protein labeled with different dyes also allows to discern and characterize interactions between protein monomers.

1.5. PHOTON STATISTICS

One of the thorniest issues in studying biomolecular complexation is that of stoichiometry. While measurement of the rotational or translational diffusion coefficient using fluorescence anisotropy, analytical ultracentrifugation, light scattering, or even correlation spectroscopy can sometimes provide limits for the stoichiometry, or its value under certain limiting concentration conditions, the interpretation of data from any of these approaches is complicated by issues of molecular shape, heterogeneous populations, and hydration considerations. In fluctuation spectroscopy however, one has the added advantage that it is a particle-counting technique, which allows for the measure of molecular brightness. Molecular brightness corresponds to the number of counts per second per molecule (cpspm) observed from the fluorescent species. One can imagine for example (Figure 1.4) that a protein–DNA complex containing a fluorescently labeled protein (1 dye molecule/monomer) emits twice as many photons as it moves through the excitation or

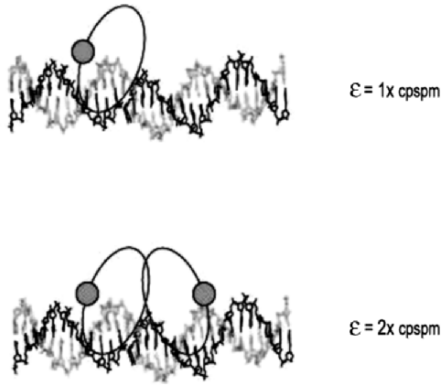

Figure 1.4. Schematic representation of a molecular brightness experiment. A protein, labeled with a green fluorophore binds to a target DNA sequence. In order to determine whether the protein binds as a monomer or a dimer, the fluorescent protein (with a 1:1 fluorophore to monomer labeling ratio) and DNA are observed at various concentration ratios (at least tenfold above the Kd of the interaction). If the complex formed is a 1:1 monomer to DNA complex, then the molecular brightness will be equal to that of the protein monomer, measured separately. If on the other hand, the complex is a 2:1 monomer to DNA complex, then the molecular brightness will be twice that of the protein monomer.

observation volume in an FCS experiment, if the protein binds as a dimer, rather than as a monomer. Thus, in addition to allowing for the determination of affinity, which is often more easily attained by other techniques, at least *in vitro*, fluctuation spectroscopy can offer a rather unambiguous measure of complex stoichiometry. Moreover, particle counting (Berland *et al*., 1996), molecular brightness (Chen *et al*., 1999b; Margeat *et al*., 2001), or distribution analysis (Kask *et al*., 2000; Palo *et al*., 2000) (discussed subsequently), all of which are capabilities of fluorescence fluctuation measurements, may be more appropriate for monitoring complexation events involving small relative changes in molecular weight.

While the autocorrelation function describes the temporal fluctuation of the fluorescence signal, several other methods rely on the analysis of the amplitude of these fluctuations. These analysis procedures allow the extraction from the fluorescence fluctuation data of the average number $\bar{N}$ and the molecular (or specific) brightness ε of the molecule in the observation volume, related to each other through the relation

$$\varepsilon = \frac{\langle k \rangle}{\bar{N}}, \tag{1.19}$$

where $\langle k \rangle$ is the average number of photon counts per unit of time, and ε is expressed as the mean count rate per molecule (in cpspm). It is especially useful to determine this latter parameter when, during the interaction process under study, the molecular weight of the species, and thus the diffusion coefficient, does not change dramatically, while the molecular brightness is strongly affected. This is the case for example during the dimerization of a labeled species, where the diffusion coefficient increases on average by only 26%; whereas the molecular brightness increases by 100% (if the two fluorophores do not interact). It is important to remember that in order to compare the values of molecular brightness, one must not vary any of the instrument parameters. That is to say that the focus, the excitation intensity, and the detection efficiency must remain constant.

We review here the different methods that have been developed to extract from the fluorescence fluctuation data either the molecular brightness alone or the molecular brightness and the diffusion coefficient at the same time, thus allowing to resolve complex mixtures of biomolecules based on these different parameters. The theories presented here take into account in various ways the geometry of the observation volume and the effects of shot noise that arises from the randomness of the fluorescence emission and detection processes. We limit ourselves to the analysis of the fluctuation data in the ''low ensemble'' concentration regime, where several molecules can be present in the observation volume. For the ''single molecule'' regime, i.e., when there is always less than 1 molecule at the time in the observation volume (i.e., typically when the fluorophore concentration is less than 0.2 nM), the methods presented here are still valid with few adjustments, but it becomes possible to analyze individually the signal from each single molecule, which appears as a ''burst'' of fluorescence above the background (see Deniz *et al*., 1999 and references therein).

1.5.1. Moment Analysis

This analysis method, introduced by Qian and Elson (Qian and Elson, 1990a,b), uses the first order moment $\langle k \rangle$ (average) and the second order moment $\langle \Delta k^2 \rangle$ (variance) of the photon counts to calculate the fluctuation amplitude, $G(0)$. This $G(0)$ value, which depends on the average number of molecules in the excitation volume $\bar{N}$, is related to the average and the variance of the fluorescence intensity through the relation

$$G(0) = \frac{\langle \Delta F^2 \rangle}{\langle F \rangle^2} = \frac{\langle F^2 \rangle - \langle F \rangle^2}{\langle F \rangle^2}. \tag{1.20}$$

The first two moments of the fluorescence intensity can be related to the moments of the photon counts

$$\langle F \rangle = \langle k \rangle \tag{1.21}$$

and

$$\langle F^2 \rangle = \langle k^2 \rangle - \langle k \rangle. \tag{1.22}$$

Thus, $G(0)$ can be rewritten in terms of photon counts

$$G(0) = \frac{\langle \Delta k^2 \rangle - \langle k \rangle}{\langle k \rangle^2}. \tag{1.23}$$

With this method, $G(0)$ can be calculated in a fast and model-independent manner. However, in order to recover the parameters, $\bar{N}$ and ε, it is necessary to know the geometric factor γ, see Eq. (1.6) and thus the calculation is not model-independent anymore. The factor γ, which depends on the shape of the PSF, equals $1/2\sqrt{2}$ for a 3D Gaussian profile and $3/4\pi^2$ for a Gaussian–Lorentzian profile. For the case of a single species, $\bar{N}$ is given by

$$\bar{N} = \frac{\gamma}{G(0)} \tag{1.24}$$

and ε is calculated using Eq. (1.19) (Qian and Elson, 1990b).

Thus, for the simple case of a single species in solution, it is straightforward to recover its brightness and concentration in a computationally simple and rapid manner, which provides a convenient means of checking the quality of a data set. However, when more species are present in solution, it is necessary to take into account the higher-order moments of the photon counts. Although suggested as early as 1990 and applied to the detection of large fluorescent beads, this approach has not been fully explored until recently, with the introduction of the factorial cumulants method (see later).

1.5.2. PCH and FIDA: Fitting the Photon Counts Distribution

Instead of calculating the moments of the photon-counts distribution, it is possible to fit this histogram directly and thus to use more information to extract

the molecular brightness and occupancy. The various sources of fluctuation that account for the shape of the distribution have to be explicitly taken into account, i.e., the shot noise, the fluctuation in fluorescence intensity caused by the diffusion of the molecules in an inhomogeneous detection profile, and the fluctuation of the number of particles within this observation volume. Two methods have been developed quasi-simultaneously to describe these distributions, the photon-counting histogram (PCH) (Chen *et al*., 1999b) and the fluorescence intensity distribution analysis (FIDA) (Kask *et al*., 1999), which differ mainly in their approach to the treatment of the observation volume. In the original description of the PCH, the observation volume generated by the two-photon excitation process was adequately modeled with a 2D Gaussian–Lorentzian function. The parameters of the function were directly recovered from the fit, and thus no calibration is needed. In the case of one-photon excitation, however, a 3D Gaussian function fails to describe the histogram correctly, and thus additional parameters have to be included in the fit to take into account the contribution to the histogram of the photons coming from the out-of-focus regions (Huang *et al*., 2004). In the case of FIDA, also developed for one-photon excitation, the observation volume is described with a polynomial function with up to three parameters. This approach, although fast and versatile, introduces several other fitting parameters, without any specific physical meaning, and thus it is necessary to calibrate the observation volume using solutions of known dyes before the analysis of unknown samples.

We must note that one condition must be fulfilled for these types of analysis to be valid; that is, that the molecules be quasi-immobile within the observation volume during the counting time interval used to build the histogram. Practically, this means that the counting time interval has to be at least ten times smaller than the diffusion time of the molecule. If this sampling interval is too long relative to the diffusion time, the PCH/FIDA theory breaks down. Thus, for an accurate evaluation of the occupancy and the brightness of the species under study, it is necessary either to determine the diffusion time using FCS analysis or to perform a PCH/FIDA analysis at various counting time intervals, and check that the recovered parameters are constant.

All these methods have the ability to extract from the histograms the molecular brightness and concentration of a mixture of species with different brightness, provided that the signal statistics are sufficient (Muller *et al*., 2000; Huang *et al*., 2004). They have been successfully applied to study ligand–protein (Chen *et al*., 2000; Rudiger *et al*., 2001; Scheel *et al*., 2001), protein–protein (Margeat *et al*., 2001), or polymer–oligonucleotides interactions (Van *et al*., 2001).

1.5.3. Fluorescence Intensity Multiple Distribution Analysis

Since temporal fluctuation analysis (i.e., FCS) and amplitude fluctuation analysis (PCH, FIDA, and others) are performed on the same original data set, it is possible in principle to combine these two types of analyses to extract simultaneously, and with a better accuracy, the diffusion time, concentration, and brightness parameters of the various species present in solution.

In the FIMDA, photon-counting histograms are built for various lengths of counting intervals (binning times), and analyzed simultaneously (Palo *et al*., 2000). We remind the reader that for the analysis of the photon-count distributions presented earlier, one fundamental assumption was that the binning time is chosen to be short enough, so that the molecules can be safely assumed to be immobile in the observation volume during the integration time. However, for the FIMDA analysis, this hypothesis is not valid anymore, and thus, for each binning time, only an *apparent* brightness is recovered; then, the dependence of the variation of this apparent brightness on the width of the time window allows an estimation of the diffusion time. Although this estimation is indirect, as compared with the direct fitting of the correlation function $G(t)$ performed in an FCS analysis, it has been shown that the accuracy of the recovered diffusion time by both methods is equivalent (Palo *et al*., 2000).

1.5.4. Fluorescence Cumulant Analysis

Recently, a new theory called Fluorescence Cumulant Analysis (FCA) has been developed and tested to extract directly without any fitting, the concentration and brightness parameters of several species, by exploiting the factorial cumulants of the photon counts (Muller, 2004). The cumulants are related to the moments of the photon counts, but are more convenient from a mathematical point of view. Indeed, cumulants of the sum of statistically independent variables are simply given by the sum of the cumulants of the individual variables. Since two cumulants are necessary to determine the brightness and concentration for one species, $2n$ statistically significant cumulants (i.e., statistically significant cumulants up to the $2n$th order) will be necessary to resolve a mixture of n species (Muller, 2004). An error analysis method was developed for the factorial photon-count cumulants, allowing the evaluation of the relative error of each experimental cumulant and thus its statistical significance. Knowing the number of significant cumulants, it is straightforward to determine how many species can be resolved. If necessary, the acquisition time can be increased to obtain the $2n$ statistically significant cumulants necessary to resolve the n species. An extension of this theory, called time-integrated fluorescence cumulant analysis (TIFCA), allows one to recover the diffusion time of the diffusing species in addition to their brightness values and concentrations. This is achieved by rebinning the data taken at short sampling time, calculating the experimental cumulants of the photon counts as a function of binning time, and then comparing them to theoretical models. This analysis procedure allows not only determination of the diffusion time, but also improves considerably the accuracy of the determination of the brightness and concentration, without increasing the acquisition time (Wu and Muller, 2005).

1.5.5. Photon-Count Distribution Analysis in Multiple Channels: Dual-Color PCH and 2D-FIDA

To examine interactions involving several partners, it is often desirable to label them with spectrally distinct fluorophores. Like cross-correlation analysis, dual color

fluorescence fluctuation data can be analyzed in terms of photon-count distributions. Two similar approaches have been developed for this purpose, allowing to extract concentration and brightness parameters in each detection channel : two-dimensional fluorescence intensity distribution analysis (2D-FIDA) (Kask *et al*., 2000) and the dual-color photon counting histogram (dual-color PCH) (Chen *et al*., 2005). So far, 2D-FIDA has been mainly used in high throughput screening studies of protein–ligand interaction (Kask *et al*., 2000; Schilb *et al*., 2004). The introduction of dual-color PCH allows us to resolve CFP/YFP mixtures *in vitro*, based on their relative brightness in each detection channel, an usually difficult task due to the high spectral overlap between these two probes (Chen *et al*., 2005).

1.5.6. Monitoring Diffusion Time, Brightness, Concentration, and Dual-Color Coincidence with Photon Arrival-Time Interval Distribution

In order to analyze the photon streams in one or two channels in the ''single molecule'' and ''low ensemble'' regime, Laurence *et al*. took a different approach than analyzing the photon-counts distribution binned in evenly spaced intervals (Laurence *et al*., 2004). Instead, in their method called photon arrival-time interval distribution (PAID), a 2D histogram is built, based on the observation of time interval between photons, which emphasizes ''photon-rich'' time intervals, where molecules are present in the detection volume. The *x*-axis is the time interval between two photons (the *start* and *stop* photons, not necessarily consecutive), while the *y*-axis represents the number of photons in the *monitor* channel counted in the time interval between the *start* and *stop* photons. For a one-color experiment, the *start*, *stop*, and *monitor* channels are the same, while in a two-color experiment, each channel can be ''yellow'' or ''red,'' and thus eight different PAID histograms are built with the different combinations. Interestingly, a simple collapse of the histogram on the time axis provides the FCS correlation function. The PAID histogram is fitted using a model that includes a numerically approximated, non-Gaussian detection volume. The various parameters characterizing the diffusing species are recovered with an accuracy comparable to FIMDA. However, PAID is also suitable for dual-color experiments, and thus provides a unique tool for the sorting of a heterogeneous mixture in the ''small ensemble'' regime. As a demonstration, a mixture of various components of transcription complexes has been successfully resolved, showing the presence of : (1) free DNA (labeled with a single red dye, Cy5), (2) species associated to σ^{70}, the initiation factor (labeled with a single yellow dye, Tetramethylrhodamine), (3) DNA–RNA Polymerase-σ^{70} complexes (labeled with a single yellow and a single red dye), and (4) σ^{70}-aggregates (labeled with several yellow dyes) (Laurence *et al*., 2004).

As shown previously, a large number of methods that allow the analysis of the same sets of fluorescence fluctuation data have been developed, to extract the brightness, concentration, and diffusion time of the species present in solution. Most of them are relatively recent, and have not been applied to solve problems of biological

interest, except for PCH and FIDA. These two methods are the simplest to implement and are already included in the software of commercial FCS systems. However, great care has to be taken in the interpretation of the data, and it is important to be able to check that the data set used contains a sufficient amount of information to resolve the various species in solution without ambiguity (Muller *et al*., 2000). From this point of view, the TIFCA method, which facilitates the calculation of the statistical significance of each of the factorial photon-count cumulants, seems to be the most advanced approach to avoid any ambiguity (Wu and Muller, 2005).

1.6. EXAMPLES OF THE USE OF FLUCTUATION SPECTROSCOPY TO STUDY BIOMOLECULAR INTERACTIONS

1.6.1. Resolution in the Measurement of Biomolecular Diffusion Times

Even with a well-characterized FCS setup and without inclusion of other photophysical processes (i.e., triplet-state emission, blinking, etc.) fitting correlation functions measured from a generic sample with the expression in Eq. (1.3) is not generally straightforward as possible polydispersity in size (D) and brightness (ε) of the fluorescent particles introduces a multiparameter fit that is difficult to handle (see Starchev *et al*., 1999; Chen *et al*., 1999a; Krichevsky and Bonnet, 2002, for discussion). Indeed, as in dynamic light scattering and similar scattering techniques, extracting the distribution of sizes from the measured FCS correlation functions requires solving a mathematical inverse problem, a challenging task since the problem is ill-posed (Meseth *et al*., 1999; Starchev *et al*., 1999). For example, a group of diffusing particles with three distinctly different sizes might yield a correlation function that is indistinguishable from that deriving from a continuous distribution spanning the same size range. Moreover, for a complete and consistent fitting of the FCS correlations, *a priori* knowledge of the distribution of brightness of the diffusing fluorescent particles is required. This information cannot be derived from the correlation functions, but rather must be obtained separately from other measurements such as photon histograms (Chen *et al*., 1999a; Krichevsky and Bonnet, 2002). These requirements impose some limitations on the technique, as shown by Meseth *et al*. (1999), who demonstrated that the resolution limit of FCS depends on several factors including difference in size between particles as well as their concentration and brightness.

In Figure 1.2 we include normalized correlation function measured from solutions of bovine serum albumin (BSA) (66 kDa), αβ–tubulin dimers (100 kDa), and 8-tubulin rings (∼800 kDa) in PBS or PIPES buffer. All functions in Figure 1.2 were collected at room temperature. Note the uniform time shift of the correlation functions with increasing molecular weight of the particles, demonstrating first the time-resolution of the FCS setup. Though not shown in Figure 1.2, the measurements on TAMRA and Alexa 546 were extended to 0.1 μs at high-count