Rapid characterization of green fluorescent protein fusion proteins on the molecular and cellular level by fluorescence correlation microscopy

ROLAND BROCK*, GYÖRGY VÁMOSI†, GYÖRGY VERÉB‡, AND THOMAS M. JOVIN§

Department of Molecular Biology, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, D-37077 Göttingen, Germany

Communicated by Manfred Eigen, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany

ABSTRACT

Fluorescence correlation microscopy (FCM) was applied to characterize fusion proteins of the green fluorescent protein (GFP) on the cellular as well as molecular level within seconds in an integrated instrument. FCM combines the inherent sensitivity and high spatial resolution of fluorescence correlation spectroscopy with fluorescence imaging and micropositioning, thereby providing a spectrum of molecular information in the cellular context. Signatures of characteristic parameters derived from the autocorrelation functions served to distinguish a GFP fusion protein of the epidermal growth factor receptor (EGFR) from GFP fluorescence in the endoplasmic reticulum and cytoplasm. Diffusion constants measured for free transiently expressed GFP reproducibility values reported previously with other techniques. The accessible concentration range extends from millions to only a few thousand molecules per cell, with single molecule detectability in the femtoliter detection volume. The detailed molecular characterization offered by FCM is fully compatible with automation in sample identification and detection, offering new possibilities for highly integrated high-throughput screening.

New techniques are constantly evolving for high-throughput screening (HTS) assays in cellular systems as well as in vitro (1). At present, however, analytical approaches are mutually exclusive, providing information either on the level of the whole cell or on the molecular level with isolated target molecules. In the functional characterization of genomic sequence information and the exploitation of this information for drug screening, experimental strategies integrating both levels of complexity are highly desirable. In vitro screening with isolated potential drug targets carries the risk of introducing artifacts due to exclusion of the physiological cellular context. Furthermore, proteins representing important clinical intervention points in cancer therapy, such as transmembrane receptors, are difficult to generate for use in in vitro model systems. Cellular screens based on reporter readouts alone, however, fail to confirm unequivocally that a specific molecular interaction or mechanism has been targeted. The green fluorescent protein (GFP; ref. 2) of the jellyfish Aequorea victoria can be fused to intracellular proteins. The use of GFP as a fluorescent tag for intracellular protein trafficking and dynamics in vivo has alleviated the requirement for protein purification, external fluorescence labeling, and microinjection, which has hitherto limited the number of accessible experimental systems. However, in vivo HTS using fluorescent techniques will require sensitive and versatile screening technology. Once a potential target is identified and characterized, one would wish to proceed to drug screening on the same instrument.

Fluorescence correlation microscopy (FCM; refs. 3 and 4) has recently been introduced as an integrated experimental approach combining the sensitivity and spatial resolution of confocal fluorescence correlation spectroscopy (FCS; refs. 5 and 6) with high-sensitivity imaging, three-dimensional micropositioning, and micromanipulation in live-cell microscopy. Confocal FCS measures fluctuations in the number of fluorescent molecules diffusing and/or reacting within femtoliter volumes and occurring over a very large temporal range (microseconds to seconds). The method is unique in providing a determination of the absolute number (i.e., concentration) of molecules and their interactions, matching the challenges outlined above. Applications of FCS with immediate relevance to pharmaceutical lead searches include the measurement of diffusion, aggregation, photophysical characteristics, receptor-ligand interactions, DNA hybridization, and enzymatic reactions (7) both at the single-molecule level and in HTS applications (8).

To demonstrate the applicability of FCM in cell biological screening, we analyzed the diffusion of a fusion of GFP with the human epidermal growth factor receptor (EGFR) and defined the nature of GFP fluorescence localized to the ER and cytoplasm in subpopulations of cells, using for comparison transiently expressed free GFP. The EGFR is a prototypic member of the large family of transmembrane growth factor receptors with tyrosine kinase activity (9). Ligand binding leads to activation of the kinase and autophosphorylation of tyrosine residues in the intracellular C-terminal segment. Signal transduction proceeds through binding of cytoplasmic proteins via SH2 and PTB domains (10, 11) to these phosphotyrosines. Screening tasks in such systems demand a rapid quantitation of proteins localized in both cellular compartments, the plasma membrane, and the cytoplasm. By using FCM, protocols can be devised for rapid molecular characterization, which are amenable to automation in such complex screening applications.

MATERIALS AND METHODS

Cell Culture. Chinese hamster ovary cells (obtained from Y. Yarden, Weizmann Institute, Rehovot, Israel) and transfected cell lines were grown in a 5% CO2 humidified atmosphere at 37°C in DMEM supplemented with 10% FCS, 104 units/liter

Abbreviations: FCM, fluorescence correlation microscopy; FCS, fluorescence correlation spectroscopy; GFP, green fluorescent protein; HTS, high-throughput screening; EGFR, epidermal growth factor receptor.

*Present address: Institute of Organic Chemistry, University of Tübingen, D-72076 Tübingen, Germany.
†Present address: Cell Biophysical Research Group of the Hungarian Academy of Sciences, H-4012 Debrecen, Hungary.
‡Present address: Department of Biophysics and Cell Biology, Medical University of Debrecen, P.O. Box 39, H-4012 Debrecen, Hungary.
§To whom reprint requests should be addressed. E-mail: tjovin@mpic186.mpibpc.gwdg.de.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

PNAS is available online at www.pnas.org.
penicillin G, and 100 mg/liter streptomycin sulfate. Cells were propagated every 3 to 4 days.

**Generation of EGFR-GFP Fusion Proteins and Transfected Cell Lines.** The EGFR-GFP fusion protein was generated as described elsewhere (12). The EGFR was derived from an EGFR cDNA (13) in pcDNA3 (Invitrogen; obtained from Y. Yarden) and cloned into the pEGFP-N3 plasmid (CLONTECH). For transient transfection, Chinese hamster ovary cells were seeded in 35-mm Petri dishes onto 12-mm glass coverslips at a confluency of 10-15%. The next day, DMEM was replaced with 0.8-ml OptiMEM (GIBCO/Life Technologies) supplemented with antibiotics. One microgram of the excitation focus (Cal, Brattleboro, VT) and a BP515–545 detection filter (Delta and detected with a 500DRLP dichroic mirror (Omega Optics)) were incubated separately with 100 μl of OptiMEM for 15 min and then for 30 min after mixing the two solutions at room temperature. The cells were transfected in 1 ml total volume for 24 h and used for microscopy 48 h later. Cells stably expressing the fusion construct were generated in the same way with selection in medium supplemented with 0.4 mg/ml G418 (GIBCO) and two rounds of cell sorting (Epics Elite, Coulter) at 2-wk intervals. The structural and functional integrity of the EGFR-GFP fusion protein localized to the plasma membrane was characterized in detail (12).

**FCM Measurements.** The FCM apparatus is based on a Zeiss Axiovert 35 fluorescence microscope with a C-Apochromat 40×/1.2 numerical aperture objective. The hardware specifications are given elsewhere (3). GFP fluorescence was excited at 488 nm and detected with a 500DRLP dichroic mirror (Omega Optical, Brattleboro, VT) and a BP515–545 detection filter (Delta Light & Optics, Lyngby, Denmark). Power densities in the excitation focus (~16 kW/cm² for intracellular measurements) were calculated according to Eggeling et al. (14). Autocorrelation functions were generated on-line with an ALV-5000/E dual channel autocorrelator board (ALV-Laservertriebsgesellschaft, Langen, Germany) and fitted off-line with Igor Pro (WaveMetrics, Lake Oswego, OR) to the autocorrelation function given below (Eq. 1). The triplet state lifetime τT, triplet fraction γT, the fractional contributions Φj, and the characteristic diffusion times τj of components j ≤ 3 were allowed to vary. An offset parameter was included to compensate for artifacts in the autocorrelation related to drift in the fluorescence signal (e.g., photobleaching) with time constants >1 s, a specific problem of the ALV-5000/E. The theory for cellular FCS measurements has been described in detail elsewhere (15). The relevant equation is

\[ G_{tot}(\tau) = 1 + \frac{1}{N_{tot}} \left( 1 + \frac{T_e^{1/(1 - \gamma_T)}}{1 - T} \right) \left( 1 - \frac{I_B}{I_{tot}} \right)^2 \sum_j \Phi_j \left( 1 + \frac{\tau}{\tau_j} \right)^{-1} \left[ 1 + (\omega_0^2/\omega_j^2)\gamma_j ]^{-1} \right]^{1/2} + \text{offset} \tag{1} \]

where \( \Phi_j = Y_j F_j / (\Sigma_j Y_j F_j) \), with the fluorescence efficiencies \( F_j \) and relative molar fractions \( Y_j \) of species \( j \). \( \omega_0 \) and \( \omega_j \) are the axial and radial waists defining the detection volume, \( I_{tot} \) the total detected signal, and \( I_B \) the intensity of uncorrelated background signal (16). The diffusional relaxation constant \( B \) is related to the diffusion constant \( D \) by \( B = \omega_0^2 / 4D \), with \( \omega_0 = 0.3 \) μm in our experimental configuration. Cells grown on 12-mm coverslips were mounted in an incubation chamber on the microscope stage in HBS (10 mM Na-Hepes, pH 7.4/135 mM NaCl/5 mM KCl/1 mM MgCl₂/1.8 mM CaCl₂/1% glucose/0.1% BSA) and used at room temperature.

**RESULTS**

**FCM-Based HTS Protocol in the Characterization of GFP Fusion Proteins.** The protocol for a high-throughput characterization of GFP fusion proteins is illustrated in Fig. 1. The experiments presented here proceeded in the same way, except that the spectrum of cells was chosen from one heterogeneous population expressing an EGFR-GFP fusion protein localized at the plasma membrane (Fig. 2A) or GFP fluorescence in the ER (Fig. 2B) or cytoplasm (Fig. 2C) rather than from homogeneous populations of cells expressing different fusion constructs. To enable the fast analysis of rapidly diffusing molecules, the excitation intensity was set relatively high. The signal-to-noise ratio in autocorrelation measurements depends on the number of photons detected per molecule per residence time in the detection volume and is therefore a function of excitation laser power as well as of the diffusion constant (17). For the EGFR-GFP fusion protein, the autocorrelation functions were acquired after an initial phase of laser illumination during which strong photobleaching occurred, indicative of slow-moving or immobile molecules.

**FCM-Based Characterization of Intracellular GFP Fusion Proteins.** Fig. 2 represents a compilation of the information obtained for GFP fusion proteins at different locations. The three-dimensional distribution of the fusion proteins in the cells was determined through a combination of intensified whole-field fluorescence imaging and recording of profiles along the optical axis, superior in sensitivity to conventional confocal microscopy for visualizing the low concentrations of fluorophores appropriate for detection with FCS. The morphological integrity of the cells before and after the measurements was documented by bright-field microscopy (Middle), in conjunction with the whole-field epifluorescence images. For each cell type, distinct autocorrelation functions were acquired within measurement times of ~10 s (Right). The numerical analysis of the autocorrelation functions yields the number of molecules \( N \) in the observation volume and delineates the number, nature, and relative fraction of molecular components of heterogeneous systems. In the case of the EGFR-GFP fusion protein, a slow component with a relaxation time of 190 ms contributed strongly to the autocorrelation function. In contrast, for the cytoplasmic protein, a fast component with a relaxation time of 810 μs dominated the signal, consistent with a small freely diffusing protein. For the endoplasmic protein, the fast component dominated; a slow component contributed approximately one-third of the autocorrelation amplitude.

In the presence of mixed populations of fluorescent molecules in different states of aggregation as well as photobleaching, \( N \) is at best an estimate of the absolute molecule number. For the rapid and reliable discrimination of the GFP fusion proteins, further analysis concentrated on the autocorrelation times and relative fractions under a given set of experimental conditions. The detailed analysis of all molecular states requires measurements at different laser powers, taking into account those molecules lost during initial photobleaching (4).

**Statistical Evaluation of FCM Data.** For statistical validation of the results presented in Fig. 2, the relative fractions \( \Phi_j \) of fast and slow components were plotted vs. the autocorrelation constants \( \gamma_j \) for the GFP signals in a number of cells of each class (Fig. 3). The autocorrelation functions were fitted to two or three diffusion autocorrelation terms according to Eq. 1. The autocorrelation analysis was model independent in that the number of components was not presumed to correspond to known molecular species. This approach is adequate for screening tasks in which the identity or state of the target molecules is not known \( \text{a priori} \), and an automated fingerprinting based on autocorrelation parameters is the primary goal.

The effect of photobleaching was compensated by an offset term in most fits. The component with a time constant of 1.2 s in Fig. 2C exemplifies the effect of mild photobleaching on the measured autocorrelation function.

The data for all three subcellular locations were clustered in three time domains, ~10 μs, 0.1 to 5 ms, and 10 ms to a few seconds (Fig. 3). The 10-μs relaxation time is too short to be
explained by a diffusional process. An apparent triplet term was included to account for fast nondiffusional fluorescence fluctuations (18). However, additional pH-dependent photo-physical processes leading to fluctuations of GFP fluorescence in the 100-ms time range have been recently reported (19). Different but overlapping distributions for the three subcellular locations were obtained, with a 1-ms fraction being dominant for the cytoplasmic GFP and a 100-ms autocorrelation time prevalent for the membrane GFP. The averages and standard deviations of these distributions differed by less than one standard deviation (Fig. 3C). Next, a weighted mean given by $\tau_{\text{mean}} = (\sum \phi_i \tau_i)/(\sum \phi_i)$ was calculated, and the frac-
deviation (Fig. 3B and D). Table 1 lists the diffusion constants that were derived from the diffusional relaxation times as described in Materials and Methods. For all three subcellular locations, the fast diffusional process corresponded well to the diffusion constant of free intracellular GFP (see below). The small diffusion constants were similar to values reported for the diffusion of GFP fusion proteins in the endoplasmic reticulum (20) and for a fusion of GFP with the β2-adrenergic receptor (21).

**Fast Measurements of Free Cytoplasmic GFP.** The cytoplasmic GFP was identified as a low molecular-weight truncated form of the EGFR-GFP fusion protein on the basis of the nuclear entry and short relaxation time, corresponding to a diffusion constant of \(2.4 \times 10^{-7} \text{ cm}^2/\text{s}\). To further confirm this assumption and demonstrate the reproducibility of FCM in another defined system, cells were transiently transfected with free GFP, which distributes homogeneously throughout the cell. Superposable autocorrelation functions were obtained for repetitive 5-s measurements (Fig. 4A). In two-component fits, either an additional slow or additional fast process was obtained, which affected the fitted time constant for the major component (Fig. 4B). Because all pairs of values fell within one of the categories defined above, weighted averages were calculated, corresponding to a diffusion constant of \(1.7 \times 10^{-7} \text{ cm}^2/\text{s}\). They were nearly identical for all four curves (Fig. 4B), demonstrating again that the model-independent fingerprinting approach was able to confirm unequivocally the identity of the proteins.

**Concentration Range for the Detection of GFP Fusion Proteins.** The range of working concentrations is a major determinant for the applicability of an assay technology in drug screening and intracellular analyses of molecular dynamics and distribution. In the early phase of a search for compounds binding to a cellular target, micromolar concentrations are desirable, whereas the identification of highly active binders requires nanomolar concentrations. Furthermore, the technique needs to cope with a spectrum of expression levels of intracellular GFP fusion proteins. The sensitivity of FCS reaches down to the single molecule level, corresponding to subnanomolar concentrations. The upper limit is about 1,000 molecules in the detection volume, i.e., concentrations of 0.1–1 μM. The autocorrelation amplitude is inversely related to the number of molecules in the detection volume. If the fluctuation amplitudes caused by molecule diffusion are too low, the calculation of the autocorrelation functions is compromised by detection noise. Figs. 2 and 4 include examples of the mean and upper concentration ranges, and in Fig. 5 a measurement at the single-molecule limit is presented. The amplitude of the autocorrelation function \(>1\) as well as the spikes in the count trace most likely corresponded to single EGFR-GFP fusion proteins or aggregates of receptors passing through the detection volume.

**DISCUSSION**

Our results demonstrate the realization by FCM of a complex cell biological screening task in the characterization of cellular proteins, benefiting from the imaging capability, the position-accuracy, and the diminutive confocal measurement volume, as well as the versatility of GFP. GFP fusion proteins localized to different subcellular compartments were discriminated with measurement times of a few seconds on the single-cell level and with single-molecule sensitivity.

The analysis of the autocorrelation curves with a model-independent approach introduced some degree of arbitrariness. We attempted to achieve fits lacking systematic errors by varying the starting conditions. The reproducibility of the fitted values was high. After binning the values within defined intervals, the parameters of plasma membrane EGFR-GFP differed from those for the other two compartments by more
than one standard deviation. HTS with FCS based on overlapping distributions of characteristic parameters for in vitro systems has recently been presented (8). A model-independent data analysis of the kind featured here will be even more appropriate if unknown proteins, e.g., from a GFP fusion cDNA expression library, are characterized.

We note that the diffusion constant derived from the long relaxation constant of the EGFR determined by FCM in the present experiments was greater by about one order of magnitude than the values obtained in previous measurements using other techniques (22, 23). This result may be attributable at least in part to the apparent reduction of diffusion constants in the presence of photobleaching (24), especially at the higher laser powers selected for fast measurements of freely diffusing intracellular molecules. However, the experiment in Fig. 5 conducted at one-fourth of the laser power and others carried out at still lower intensities were consistent with the results of Fig. 2. Particular screening tasks or detailed molecular characterizations of slowly diffusing molecules will require adjustment of this critical parameter. Such a fast diffusing component has not been reported so far for an endoplasmic protein or for a transmembrane receptor. Whereas Fluorescence Recovery After Photobleaching (FRAP) detects only recovery of fluorescence by diffusion into the detection volume, FCS is sensitive as well to short-distance movements occurring within the detection volume. Limited proteolytic degradation and release of free GFP into the cytoplasm may also occur. Control experiments with untransfected Chinese hamster ovary cells excluded autofluorescence as a source for the observed autocorrelations (15). Further experiments will be required to define the nature of the fast process in detail.

The work presented here demonstrates the analysis of GFP fusion proteins on the cellular as well as molecular level with FCM. A very similar protocol applies to in vivo drug screens. EGFR and other receptor tyrosine kinases are overexpressed in a number of tumors (25, 26). Because signal transduction proceeds through the interaction of proteins with the phosphotyrosine motifs of these receptors, these receptors and their binding partners have been designated as potential drug candidates.

**Table 1.** Diffusion constants of cellular GFP fusion proteins

<table>
<thead>
<tr>
<th>Membrane</th>
<th>ER</th>
<th>Cytoplasm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>D_{average} [cm^2/s], %</td>
</tr>
<tr>
<td>Fast</td>
<td>2.1 ± 0.8 × 10^{-7} 35 ± 7</td>
<td>1.2 ± 0.4 × 10^{-7} 62 ± 9</td>
</tr>
<tr>
<td>Slow</td>
<td>1.6 ± 0.9 × 10^{-9} 51 ± 14</td>
<td>2.9 ± 1.5 × 10^{-9} 29 ± 5</td>
</tr>
</tbody>
</table>

Given are the average diffusion constants and relative fractions for GFP fusion proteins derived from the autocorrelation times (1-ms and 100-ms processes in Fig. 3D). The fast components correspond well to the diffusion constant of 1.7 ± 0.4 × 10^{-7} cm^2/s for free intracellular GFP. The standard deviations were derived from all the measurements shown in Fig. 3B.

**Fig. 4.** Intracellular autocorrelation functions (A) and fits (B) of repeated 5-s measurements of transiently expressed free GFP in the cytoplasm. The ordinate to the right in B corresponds to the relative fractions \( \phi \) of the components in the fits. A slow (\( \times \)) or fast (\( \ast \)) process was obtained in addition to the major component; the arrow points at the weighted means. The average diffusion constant of cytoplasmic GFP was 1.7 ± 0.4 × 10^{-7} cm^2/s, in good agreement with a value reported previously (29).

**Fig. 5.** FCM at low levels of intracellular EGFR-GFP fusion proteins. (A) Fluorescence count trace. The peaks in the fluorescence count trace likely represent single EGFR-GFP molecules diffusing through the detection volume. (B) Autocorrelation function and parameters derived from a two-component fit. The number of molecules derived from the autocorrelation amplitude was corrected for the uncorrelated background fluorescence, derived from peak-free count intervals according to Eq. 1. The excitation intensity was 3.8 kW/cm^2.
targets (26, 27). The discrimination between a free cytoplasmic and a membrane protein presented here will be essential for identifying drugs that inhibit these interactions in vivo. We envisage screening for potential inhibitors of the binding of a cytoplasmic protein to a transmembrane receptor based on changes in the autocorrelation function as well as in the distribution of the proteins in the fluorescence images, thereby integrating the molecular and cellular levels in the lead search. The concentration range of three orders of magnitude will enable the acquisition of dose-response curves of compounds with different biological activities. FCM is compatible with expression of standard cytomegalovirus promoter-driven GFP fusion constructs. However, regulatable expression (28) represents the next step for better control of experimental conditions. FCM will enable not only analyses at low concentrations of GFP fusion proteins, thereby reducing the risk of artifacts because of protein overexpression, but also the quantitation of the number of these probe molecules inside the cell.

With short measurement times, high sensitivity, and automation achieved by a combination of image-based object recognition and motorized high-precision positioning, FCM constitutes a unique tool for meeting the challenge of intracellular HTS.

R. B. received a fellowship from the Studienstiftung des deutschen Volkes. The cooperation between T.M.J., G. Vamosi, and G. Vereb was supported by Deutscher Akademischer Austauschdienst-Hungarian Fellowship Commission joint grant for exchange of researchers no. 91/1997. G. Vereb was supported by Orszoś Tudományos Kutatsí Alap (Hungary) grant F025210 and the Alexander von Humboldt Foundation. We thank Didier Marguet, Johns Hopkins University, for plasmids and Rolando Riviera-Pomar for critical reading of the manuscript.