Single molecule methods with applications in living cells
Fredrik Persson, Irmeli Barkefors and Johan Elf

Our knowledge about dynamic processes in biological cells systems has been obtained roughly on two levels of detail; molecular level experiments with purified components in test tubes and system wide experiments with indirect readouts in living cells. However, with the development of single molecule methods for application in living cells, this partition has started to dissolve. It is now possible to perform detailed biophysical experiments at high temporal resolution and to directly observe processes at the level of molecules in living cells. In this review we present single molecule methods that can easily be implemented by readers interested to venture into this exciting and expanding field. We also review some recent studies where single molecule methods have been used successfully to answer biological questions as well as some of the most common pitfalls associated with these methods.

Addresses
Department of Cell- and Molecular Biology, Science for Life Laboratory, Uppsala University, Sweden

Corresponding author: Elf, Johan (johan.elf@icm.uu.se)

Current Opinion in Biotechnology 2013, 24:xx–yy
This review comes from a themed issue on Systems biology
Edited by Orkun S Soyer and Peter S Swain

0958-1669/$ – see front matter, Published by Elsevier Ltd.
http://dx.doi.org/10.1016/j.copbio.2013.03.013

Introduction
There are two main reasons for the use of single molecule techniques in living cells. First, it enables studies of intracellular processes that involve scarce molecules in individual cells. Second, it makes it possible to gather information about molecular kinetics that traditionally only could be studied in vitro using biochemical assays with purified components allowing for rapid mixing and quenching. In vivo, such synchronized perturbations from the equilibrium are usually not possible, creating a need for single molecule techniques in the study of intermolecular binding and dissociation reactions in the living cell. Straightforward single molecule techniques based on changes in diffusion properties or co-localization have therefore proven useful as in vivo complements to traditional biochemical techniques for studying intermolecular interactions. Using more advanced techniques, based on, for example, single molecule FRET, it is also possible to gain access to detailed information about intramolecular states [3]. However, even simple single molecule studies in living cells do present specific challenges and a successful experiment relies on bright and specific labels as well as low fluorescence background under reproducible experimental conditions in immobilized healthy cells. We will briefly review some of the recent attempts to overcome these problems.

Single molecule fluorescence microscopy
The most common principle for single molecule detection is regular wide field fluorescence microscopy. The requirements on the system, as compared to standard fluorescence microscopy, are dictated by the low number of photons emitted from a single molecule. The most crucial parts of the set-up are therefore: a highly photon sensitive camera (usually based on EMCCD technology), a high numerical aperture objective, high quality filters, and a stable light source [1*,2].

The single molecule epithet essentially comprises that only a few molecules are fluorescent (or rather detected) simultaneously. Many macromolecules, for example, some transcription factors, are so dilute that any diffraction limited fluorescence microscopy experiment becomes inherently single molecule [4*]. For more abundant proteins there are methods, for example, photoprotection, to ensure that only a fraction of the molecules are fluorescent at any given time [5,6**].

Labeling
There are two major classes of fluorescent labels: synthetic dyes [7] and fluorescent proteins (FPs) [8]. Synthetic dyes have the benefit of superior photo-physical properties and are more robust than FPs; however, labeling molecules inside living cells generally introduces some non-negligible issues, for example, cell wall permeability, target specificity, and labeling stoichiometry. These issues often result in insufficient labeling or excessive dye that contributes to background noise and cross reactivity.

Considering these challenges it is not hard to understand why synthetic dyes have been used mainly to study membrane proteins in living cells. However, recently developed methods based on fusion protein tags [9,10] showed increased usability for live-cell imaging [11*,12†].

In contrast, genetically encoded FPs naturally achieve perfect stoichiometry as well as target specificity. However, the use of FPs introduces the potential problem of maturation, which determines how fast the fluorophore can be detected after expression.
2 Systems biology

Furthermore, FP-labeling will often interfere with protein functionality and can also cause protein aggregation [13**,14].

Proteins are the most commonly tagged biomolecules; however, using hybridization probes [15] it is also possible to tag for example single RNAs in *vivo*. Alternatively, a two-step labeling method can be used which is based on an introduced loop structure in the RNA, potentiating binding to a fluorescently tagged MS2 coat protein [16].

If the target of interest is abundant such that it is impossible to separate the molecules if they are simultaneously excited, there is a need for fluorophores that can be switched on and off, or that can be converted between different fluorescent states (e.g. red and green). These fluorophores are often referred to as photo-switchable, photo-activatable or photo-convertible depending on their detailed mechanism and exist in a multitude of variants, often as derivatives of well-known FPs [5,8]. The use of photo-convertible dyes requires that an additional light source can be introduced in the light path in the optical set-up (Figure 1a).

**Suppressing the auto fluorescent background**

One of the major challenges of fluorescent microscopy in the living cell is the high background caused by cellular auto-fluorescence and fluorescent species in the culture medium, for example, Riboflavin. This is also one of the major reasons why many pioneering studies have been performed in bacteria under conditions with low auto-fluorescence [17,18]. Since cellular auto-fluorescence is usually largest in the green part of the spectrum, it is wise to favor red fluorophores if other desired properties, for example, maturation rate, are sufficient [19]. Another efficient way of reducing background is to selectively excite the fluorophores in a well-defined plane. This can be implemented either by side-ways illumination with a focused beam (SPIM), or by slightly shifting the excitation angle in a normal microscope to cause total (or near total) internal reflection (TIRF). Recent advances in SPIM enables excitation of very thin planes, but this usually requires non-standard optical set-ups [20]. As it is easy to implement, TIRF is by far the most widely used technique to minimize background fluorescence. As the name implies the light used for TIRF imaging never penetrates into the sample; instead the fluorophores are excited by an evanescent field that forms at the glass-sample interface. Since the evanescent field decays exponentially, only molecules that are very close to the interface (within 100–300 nm) will be detected, making TIRF optimal for imaging of membrane proteins. However, conclusions regarding structures or processes in other parts of the cell should be made with caution.

In a variation of TIRF, entitled highly inclined and laminated optical sheet (HILO) microscopy, a thin sheet of excitation light penetrates the sample at an angle, making the method in a way analogous to SPIM [21]. HILO thus has the benefit of being able to detect fluorophores inside the cell.

**Localization**

When the target molecule has been adequately labeled we need ways to identify the single fluorophore and to localize it with high accuracy (20 nm). Since the accuracy by which a single photon is transmitted through the optical system is limited by diffraction to $\sim \lambda / 2 \approx 250$ nm, the average localization of several photons ($N$) is typically used to estimate the location of the point source to accuracy $\sim \lambda / 2 \sqrt{N}$ [22*]. This is often achieved by fitting the detected fluorescent signal to a model point-spread function (PSF), usually a Gaussian (Figure 1c). Other approaches, such as wavelet analysis [23], stable wave detection [24] and radial symmetry [25], have also proven very powerful and recent advances have been made in separating adjacent fluorescent objects, decreasing the limitation to very sparse fluorescent emitters [26–29]. Addition of a second fluorophore makes it possible to compare the localization of different targets and to find colocalization. This approach was employed by Ptasin *et al.* to show that ParB-stimulated ParA depolymerization contributes to chromosome segregation in *Caulobacter crescentus* [30*].

Localization is not limited to planes but can easily be expanded to three spatial dimensions (3D). Traditionally 3D images have been obtained with scanning confocal, or two-photon, microscopes, which are expensive and relatively slow to operate. However, by introducing astigmatism in the optical system [31] or by using multiple focal planes [32] it is possible to obtain 3D localization information with a regular fluorescent microscope. Similar to confocal microscopy, the performance is decreased in the third dimension, usually by a factor 1.5–2. More advanced methods with higher 3D accuracy are available, including 4Pi microscopy [33], interferometric microscopy [34] and double helix PSF microscopy [35,36], but they are generally more complicated to implement.

**Application: single molecule experiments for low copy number experiments**

Many central biological processes involve molecules that are present in very low copy-numbers per cell. Relevant *in vivo* investigation of these processes requires single molecule sensitivity since overexpression of labeled targets would alter the process of interest, for example, by saturating relevant binding sites. Furthermore, when monitoring individual components in these processes their single molecule nature can be directly assessed. In this way, Lia *et al.* [37] successfully monitored polymerase exchange during bacterial DNA replication (Figure 2a) and Hammar *et al.* [4*] demonstrated the mechanisms by which transcription factors find and bind
individual binding sites in the bacterial chromosome (Figure 2b). In both these experiments single molecule detection was used to probe molecular interactions in the natural low copy number regime and not as a way to get around the problem that averaging properties over many molecules in different states of binding hides the relevant information.

**Application: single molecule experiments to study kinetics in asynchronous molecules**

The most important reason for exploring single molecule techniques in vivo might be the possibility to learn something about kinetics from an ensemble of molecules near equilibrium. If you could follow individual molecules through distinct binding and dissociation events...
4 Systems biology

Figure 2

Low copy number experiments. (a) By fluorescently tagging DnaQ (a subunit of DNA polymerase III) with YPet and observing the fluorescence intensity of diffraction limited spots, Lia et al. managed to draw conclusions regarding the polymerase exchange mechanisms in active bacterial replisomes [37]. (b) By monitoring the average time for binding of fluorescently tagged lacI transcription factors to individual operator sites, in conjunction with alterations of the chromosomal context surrounding the binding site, Hammar et al. could elucidate the mechanism of finding and binding to individual operator sites [42].

You could estimate the rates for binding and dissociation from the average times spent in different states of diffusion. Given that you can observe multiple events, you can even determine the waiting time distribution and learn more about the underlying processes. One may argue that the same information can be obtained by a FRAP experiment [38] or FCS [39]. However, since these methods do not keep track of the identity of individual molecules, the interpretation will be highly model dependent and it is hard to disentangle diffusion state occupancies from diffusion rates.

Biologically relevant dynamics occur on time scales ranging from biomolecular reactions (ms) to the life time of cells (h). To be able to study freely diffusing molecules and macroscopic reaction dynamics, a corresponding temporal resolution is necessary. If one of the reaction partners is relatively immobile, one approach is to image the mobile molecules at different time scales and determine how short exposure is required to freeze the molecules in space. The time required will give you an indication of the duration of interaction with the immobile partner.

Another way of accessing dynamic information in living cells is to perform sequential localizations of individual molecules, that is, single particle tracking (SPT). If the molecules change state of diffusion upon binding or dissociation the transition rates between different states as well as the state life-time distribution can be extracted from the trajectories. If the labeled species are abundant it is possible to use photo-activatable fluorophores to acquire many (100–10 000) independent trajectories from the same cell [40,62]. Historically, SPT has primarily been used to study diffusion of membrane proteins labeled with organic fluorophores, but recent years have
Figure 3

Kinetics in asynchronous molecules. (a) By tracking SH2 modules using TIRF microscopy and photo-convertible dyes, Oh et al. have investigated the kinetic properties of receptor tyrosine kinase signaling in vivo. To estimate the effective diffusion constants for different SH2 modules, MSD curves were calculated from single particle traces (Left) and the \( D_{\text{eff}} \) values were correlated to the dwell time of the molecules at the plasma membrane (Middle). Variability in \( D_{\text{eff}} \) for different modules favors a model with fast recombination after dissociation (Right) [47]. (b) RNA helper protein Hfq mediates post-transcriptional gene regulation by facilitating interactions between mRNA and non-coding small RNA. Persson et al. have used photo-convertible proteins to track single hmq molecules and assign them to different kinetic states based on their diffusive properties. The nature of the different states (Right) can be deduced by comparing un-treated cell (Left) to cells treated with the transcription blocking drug rifampicin (Middle) [45**].

seen an increase of SPT studies also in the cytoplasm of living cells (Figure 3) [41*,42–44,45**,46,47].

How to interpret SPT data
The main challenges of combining single localization events into trajectories are caused by signal disappearance due to, for example, blinking, exits from the focal plane, detection failure or merging and splitting events. These problems can be reduced by imaging molecules one at the time and by appropriate choice of fluorophores. For bright fluorophores, there are several methods available to recover trajectories suffering from temporal signal disappearance or merging or splitting events (Figure 1b) [48,49]. However, for many low signal applications it might be better to adopt a simpler and more restrictive approach by aborting incomplete trajectories to avoid misclassification.

Single molecule tracking data often contains vast amounts of information including different diffusive states, possibly representing complex formation or different modes of diffusion, transition rates between these states, and spatial localization. The challenge lies in extracting the information from the highly fragmented data often obtained from SPT in vivo. Traditionally Mean Squared Displacement (MSD) [47,50,51] and/or Cumulative Distribution Function (CDF) analysis [52] have been used to analyze this kind of data.

The MSD describes how far the molecules move on different time scales and the CDF describes the distribution of step lengths on a fixed time scale. Both methods can be used to identify parameters in presumed underlying models, or even to discard too simple models. The problem with both methods is that introduction of an
6 Systems biology

additional state in the model always leads to a better fit of the data which makes it difficult to draw conclusions regarding the number of ‘true’ underlying states, or to find transitions between states that are overlapping. These problems can be addressed using Bayesian statistics (BS), which has an inherent ability to handle large complex and noisy data sets (Figure 1d) [53–55]. While classical statistics usually calculates the probability of obtaining the data from a given model, BS infers the probability of a model given the data, including a priori knowledge or expectations of the model, if applicable. The main drawback is that it is quite computationally demanding.

Several recent studies use BS, for example, to localize fluorophores [27,28], link trajectories [49], deduce the mode of transport [51], analyze FRET data [56,57] or to find underlying states and transitions in SPT data [45**].

Conclusions

It should be noted that data from live cell single molecule experiments are notoriously easy to misinterpret. It is necessary to develop stringent controls on in vivo activity of labeled macromolecules and their binding properties since there are many more potential cross reactions in a living cell than in the test tube. Things that should be tested are for example that the labeled molecule can replace the wild type copy under the relevant conditions, that the binding is lost when binding sites are removed and that the labeling procedure and laser Illumination do not alter cell physiology.

In addition, reproducibility is a specific concern. Data usually have to be collected from several cells to get sufficient statistics to test quantitative models, and even if sufficient data can be obtained in one cell, several cells need to be investigated to characterize the cell-to-cell variability. In order to merge or compare data between cells, they should be investigated under similar conditions and sometimes even in the same phase of the cell cycle. Recent development in microfluidics growth chambers for live cell single molecule experiments and the corresponding image analysis software [58] will make it feasible to perform reproducible experiments in the required amount of cells. Another example of advances in automated analysis is rapidSTORM, a recently introduced open-source software for localization microscopy [59**].

Finally, single molecule experiments in living cells often need to be complemented by quantitative modeling at a similar level of detail [60] to determine if the observed data is expected given geometrical constraints and biological and experimental noise. Quantitative modeling and simulation of alternative models may also be the only way to test if these models can be discarded given the amount and quality of the available data.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest


Using a single-molecule imaging assay that examines the binding of a fluorescent LacI repressor to its target operator, Lewin and colleagues investigate the DNA-sliding mechanism and find that LacI slides a distance of 45 ± 10 base pairs on the DNA and that other DNA-binding proteins in close proximity to the operator can impede this sliding process.


By using SNAP tags with fast-switching photo-convertible dyes the authors demonstrate 2D live imaging at a spatial resolution of approximately 25 nm and a temporal resolution of 0.5 s and 3D live images at a resolution of approximately 30 nm in the lateral direction and ~50 nm in the axial direction at time resolutions of 1 s.


By combining dualcolor FPALM with direct stochastic optical reconstruction microscopy (dSTORM) Wilmes et al. manages to create triple color images to investigate the spatiotemporal organization of receptor proteins in the plasma membrane.


In this paper the authors describe an non-invasive method to assess single cell variability following cell division and highlight the problem of mislocalization of proteins due to aggregation of fluorescent labels.


A complete theoretical description of localization microscopy including a comparison between the most common statistical estimators used for fitting single emitter signals.


By employing photo-bleaching and 2D Gaussian fitting the authors perform dual color single molecule tracking to show that ParB-stimulated ParA depolymerization contributes to chromosome segregation in Caulobacter crescentus.


In this paper the authors manage to track single molecules that diffuse freely in the cytoplasm using a novel single-molecule tracking methodology.


The authors present a novel method and open source software for classifying diffusion states and extracting reaction kinetics from in vivo tracking data.


8 Systems biology


A description of a fast, easy-to-use open source software for fluorescence localization capable of handling 3D and multicolor data. In the supplementary data there is an exhaustive description of the algorithms as well as a comparison to other available software.

