Manuscript Number: COCHBI-D-14-00010R1

Title: Quantitative super-resolution microscopy: pitfalls and strategies for image analysis

Article Type: 20 Molecular Imaging (2014)

Corresponding Author: Professor Melike Lakadamyali,

Corresponding Author's Institution: ICFO-Institut de Ciencies Fotoniques

First Author: Nela Durisic, PhD

Order of Authors: Nela Durisic, PhD; Lara Laparra Cuervo; Melike Lakadamyali
Super-resolution nanoscopy is an enabling tool for biology

Multiple molecular states of super-resolution probes complicate image quantification

It’s easy to misinterpret nanoscopy images if unaware of pitfalls in data analysis

Same probe appearing many times can lead to overcounting and artificial clustering

Failed photoactivation of a probe can lead to undercounting and missed events

We describe methods for overcoming these challenges to properly quantify images
Quantitative super-resolution microscopy: pitfalls and strategies for image analysis

Nela Durisic¹, Lara Laparra Cuervo¹, Melike Lakadamyali¹*

¹ ICFO-Institut de Ciencies Fotoniques, Mediterranean Technology Park, Av. Carl Friedrich Gauss, 3, 08860, Castelldefels (Barcelona), Spain

* Correspondence should be sent to M.L.: melike.lakadamyali@icfo.es

Abstract

Super-resolution microscopy is an enabling technology that allows biologists to visualize cellular structures at nanometer length scales using far-field optics. To break the diffraction barrier, it is necessary to leverage the distinct molecular states of fluorescent probes. At the same time, the existence of these different molecular states and the photophysical properties of the fluorescent probes can complicate data quantification and interpretation. Here, we review the pitfalls in super-resolution data analysis that must be avoided for proper interpretation of images.

Introduction

The immense toolbox of fluorescence probes has enabled us to tag almost anything inside cells with high molecular specificity and in many colors. However, due to diffraction, the resolving power of an optical microscope was, until recently, limited to ~200 nm in the lateral and ~500 nm in the axial direction. This limitation obscured essential details at the nanometer length scales. Fortunately, diffraction limit has now been broken thanks to the development of super-resolution microscopy methods including stimulated emission depletion microscopy (STED) [1], saturated structured illumination microscopy (SSIM) [2], stochastic optical reconstruction microscopy (STORM) [3] and (fluorescence) photoactivated localization microscopy (PALM and fPALM) [4,5] among others. These methods are starting to enable near molecular-scale spatial resolution in biological imaging. The advent of “nanoscopy” has also brought the need for proper data analysis tools to quantify the super-resolution images. The quantitative information of interest can take several forms including characterization of image resolution, measure of nanostructure or nanocluster sizes, quantification of spatial (co-)organization of nanoclusters and “molecular counting” of protein numbers. The increased resolution imposes stringent conditions on data analysis and extra care must be taken to ensure that photophysical properties of the probes, labeling and imaging strategies used do not lead to misinterpretation of the data. As the nanoscopy field has progressed at a fast pace, several new methods have also been developed for “post-image” analysis. This review will focus on considerations one must be aware of while quantifying super-resolution images. The emphasis will be given to methods based on single molecule localization (such as STORM, PALM, fPALM) but similar points are often also important for quantifying STED images.
Basic principles

Super-resolution methods that rely on single molecule localization have been the subject of several recent reviews and for in depth information the reader is directed elsewhere [6-8]. Briefly, the image of a single molecule is referred to as the point spread function (PSF) of the microscope. Even though the size of the PSF is determined by diffraction, the molecule’s position can be precisely localized by finding the center of the PSF [9]. To avoid the problem of overlapping PSFs in a densely labeled sample, it is necessary to control the number of molecules that are fluorescently active at any given time. This active control of fluorophore density was made possible by the discovery of photoswitchable fluorophores [10-12]. In super-resolution imaging, only a small subset of fluorophores is photoactivated at any given time such that their PSFs are non-overlapping and their positions can be precisely determined. Through iterative cycles of activation and de-activation, the positions of all the fluorophores can be precisely determined, and these positions can then be used to reconstruct a high resolution image of the underlying structure.

Photoswitchable probes

A wide range of probes such as fluorescent proteins, small organic fluorophores and quantum dots are available for super-resolution imaging with single molecule localization. These probes typically fall into two categories [13,14]: (i) reversibly photoswitchable probes that can be cycled many times between bright and dark states such as small organic fluorophores (e.g. Alexa 647) [15] or photochromic fluorescent proteins (e.g. Dronpa) [16] and (ii) irreversibly photoactivatable and photoconvertible probes (e.g. PA-GFP and mEos2) [10,17]. These different categories provide complementary advantages for super-resolution imaging. Reversible probes lead to smoother images since the same structure is sampled many times [15]. However, the uncertainty in each localization makes it impossible to identify which localizations arise from the same fluorophore, especially when samples are densely labeled, thus making quantification challenging. Therefore, proper care must be taken to account for probe photophysical effects when quantifying images as described below.

Image resolution

The spatial resolution in nanoscopy can be affected by several factors including localization precision, labeling density and probe size. The localization precision mainly depends on the number of photons emitted, as well as other factors such as background noise, and pixel size [9,18,19]. Labeling density can be a major limitation to achieve high spatial resolution. Low labeling densities can cause continuous structures to appear discontinuous or clustered, resulting in a loss of detail. According to the Nyquist criterion [15,20,21], the labeling density must be such that the distance between individual localizations in the resulting image is at least half of the desired resolution.
In practice, quantifying the final image resolution can be quite difficult. While the localization precision alone can be analytically calculated [19] or experimentally determined by measuring the standard deviation of a cluster of localizations originating from a single fluorophore [3], the effects of labeling density and probe size can be challenging to determine. One strategy to quantify the overall resolution is to measure the separation distance between two closely spaced, barely resolvable structures in the image (e.g. the two walls of a hollow microtubule) [15,22]. However, finding such structures in a given image is difficult. Recently, a non-biased method based on Fourier ring correlation (FRC), originally developed for electron microscopy of single proteins [23], has been applied to super-resolution to calculate “intrinsic” image resolution directly from experimental data [24,25]. This method works by splitting the single-molecule localization data into two statistically independent sets to generate two sub-images. The Fourier transform of the two sub-images is then computed and correlated and the largest spatial frequency for which the correlation is still considerable is taken as a measure of the resolution. The inverse of this spatial frequency is the spatial resolution limit below which two objects cannot be resolved. In addition, FRC can also be used to estimate the average number of times a single emitter is localized, therefore helping to reduce over-counting artefacts (see below).

**Spatial organization and molecular counting**

In single molecule localization microscopy, super-resolution images are built molecule by molecule. Therefore, in principle, the imaging strategy should allow determining absolute numbers of molecules that exist in specific sub-cellular compartments. However, it is important to carefully examine how many molecules are missed or falsely included into the analysis.

**Artificial clustering and over-counting**

Artificial clustering and over-counting can arise due to several reasons. For example, if polyclonal primary and/or secondary antibodies are used, multiple antibodies can bind to the same protein giving rise to over-labeling. Antibodies can also physically crosslink proteins, leading to artificial clustering. Nonetheless, antibody labeling has been used to determine the relative amount and heterogeneity of synaptic proteins inside individual synapses using super-resolution imaging [26]. For determining absolute protein stoichiometry and sub-cellular protein distribution, fluorescent proteins that allow one-to-one tagging of the protein of interest may be preferable. Among these fluorescent proteins, irreversibly photoactivatable and photoconvertible ones are the best candidates for quantitative imaging, since in principle each fluorescent protein is imaged (and counted) only once. However, even the irreversible fluorescent proteins undergo transitions to non-fluorescent off states (blinking and re-activation) and can reappear multiple times during the imaging [27]. If not taken into account, this effect gives rise to over-counting and may make the image look clustered while there is no actual physical clustering of the target.
One way to account for intermittency of fluorescent protein emission is to take advantage of the time dependence of the blinking and photoactivation [28-30] (Figure 1). Typically, if the photoactivation is carried out slowly over a long time [30], blinking (or reactivation) events are closer in time compared to the photoactivation of a new fluorescent protein (Figure 1b and c). Therefore, peaks (or localizations) due to blinking can be grouped together based on a minimum dark time or cutoff time ($\tau_{\text{cutoff}}$). In this case, only localizations appearing after a time larger than $\tau_{\text{cutoff}}$ are considered to arise from a new fluorescent protein (Figure 1d). $\tau_{\text{cutoff}}$ can be determined experimentally, by imaging purified fluorescent proteins on glass. When the density of the proteins to be counted is very high, the probability of photoactivation increases and blinking and new photoactivations start overlapping in time, making counting more challenging. Therefore, fluorescent proteins with low blinking (and reactivation) rates and short blinking times are preferable (e.g. Dendra2 and PA-mCherry) [30,31]. In addition, instead of using a single value, $\tau_{\text{cutoff}}$ can be characterized as a function of protein density. An algorithm can then be used that starts with an initial value of $\tau_{\text{cutoff}}$ set to zero, counts the number of molecules and adjusts the value of $\tau_{\text{cutoff}}$ based on this number iteratively until it converges [30]. The dark time analysis has mostly been applied to super-resolution images of fluorescent proteins, since the photophysics of most small organic fluorophores (e.g. AlexaFluor647) are more complex and do not allow spatiotemporal grouping of localizations. However, recently, Zhao et al. demonstrated that the blinking events of rhodamine dyes are also temporally grouped [32]. Therefore, using genetically encoded SNAP and Halo tags to link these dyes to RNA polymerase II, they could determine the spatial organization and stoichiometry of this motor inside the nucleus.

Statistical analysis such as pair-correlation functions can also be used to correct for over-counting and artificial clustering to determine the true spatial organization of proteins and count their numbers [33-35] (Figure 2). Pair correlation function reports the probability to find a second localization a distance $r$ away from a given localization. It is given by the sum of correlations arising from multiple appearances of the same probe and the correlations due to the spatial distribution of labeled proteins. Therefore, if the first contribution is known, it can be used to correct the experimental data for over-counting and determine the true spatial distribution of proteins of interest. The correlations arising from multiple appearances of the same probe can be computed by using a calibration sample in which the probe is randomly distributed on a glass slide and imaged under the same experimental conditions (Figure 2a-c). In the absence of blinking (and re-activation) the correlation curve for the calibration sample should be flat with a value of one, consistent with random distribution of the fluorescent protein on the glass slide. Blinking and re-activation leads to higher correlations at small length scales (comparable to the localization precision), which must be corrected for in the experimental data. After this correction, the decay of the remaining correlation curve is a measure of the spatial extend of clustering (cluster size) and the amplitude of the correlation curve is related to the number of proteins within a cluster (Figure 2d-f). This type of analysis can also be applied to two-color images to determine the level of co-localization of multiple proteins. In this case, multiple appearance of the same probe does not influence the cross-correlation curve, making analysis and interpretation easier. Spatial co-organization and co-localization of proteins can also be quantified by
determining an “interaction potential” that is most likely to lead to the observed distribution of point localizations [36,37].

Finally, the FRC analysis described above [24] can also be used to extract the average number of times a molecule blinks given that the emitter activation follows a known distribution (e.g. Poisson statistics when not limited by photobleaching). For intermediate spatial frequencies, the influence of multiple localizations of the same emitter dominates the cross-correlation. If the decay of the correlation in this regime is fitted, the over-counting due to intermittency can be accounted for without the need for a calibration sample.

**Missed molecules and under-counting**

Once the effects of blinking and re-activation are corrected, the remaining number of localizations is a measure of the detected number of probes. However, many complications remain in converting this number to actual biological stoichiometry.

First, it is important that there is a one-to-one ratio between the protein of interest and the label. It is often challenging to achieve this one-to-one ratio with antibodies. While polyclonal antibodies can lead to overlabeling, the large size of the antibody can also lead to underlabeling due to steric hindrance. In practice, determining the labeling efficiency of an antibody is quite challenging and smaller probes such as aptamers and nanobodies have been demonstrated to lead to denser labeling of certain targets and higher spatial resolution in super-resolution images [38,39]. In principle, fluorescent proteins avoid this complication, giving rise to a one-to-one labeling ratio; however, the expression strategy used can cause additional problems. Transient transfection, the most common way of tagging proteins with fluorescent proteins in mammalian cells, leads to a mixture of endogenous unlabeled and over-expressed labeled protein. Ideally, one would like to determine the stoichiometry and spatial distribution of endogenous and not overexpressed proteins. With advances in genome editing, it is now possible to endogenously tag target proteins with fluorescent proteins [40,41].

Second, the imaging strategy must be such that multiple probes are not photoactivated simultaneously within a diffraction limited volume, since their images will overlap leading to missed events. The number of photoactivated probes is proportional to the power of the photoactivation laser and the total number of the remaining un-activated probes. Therefore, an imaging strategy must be followed in which the power of the photoactivation laser starts out very low and is progressively and slowly increased over time [30,42]. Similarly, the imaging period must be long enough such that all the probes are exhaustively imaged. Plotting a cumulative curve of localized molecules in each frame can give an indication of whether this condition has been satisfied [42]. The cumulative curve should increase slowly, reaching a plateau once most of the probes have been imaged.

Last but not least, failed photoactivation can lead to missed events and undercounting. A number of recent papers have used calibration standards to calculate the percentage of successful photoactivation during super-resolution imaging for a number of fluorescent proteins [31,43,44]. Durisic et al. took advantage of an *in vivo* “nanotemplate” with a well-
defined subunit stoichiometry, the human glycine receptor (GlyR) [31] (Figure 3). When expressed in Xenopus oocytes, which do not endogenously express this receptor, GlyR forms hetero-pentameric ion channels with three α- and two β-subunits [45]. The known stoichiometry made it possible to use binomial statistics to characterize fluorescent protein photoactivation efficiency by counting “steps” in single-step photobleaching or “peaks” in super-resolution intensity-time traces. Similarly, generating tandem repeats of fluorescent proteins attached to membrane proteins [44] or cytoplasmically expressing these tandem repeats at very low densities [43] allowed the use of binomial statistics to characterize the percentage of photoactivation. These results have shown that typically only 50-60% of fluorescent proteins are actually photoactivated into a fluorescent form and this number must be taken into account as a correction factor to properly count proteins in super-resolution images.

Conclusions and Outlook

Enormous amount of progress has been made in recent years in the field of super-resolution microscopy. New imaging methods with different capabilities (3D, multi-color) and new analysis methods to reconstruct super-resolution images from raw data and extract quantitative information have been developed. While the photophysical properties of the probes, labeling and imaging strategies make image analysis complex, with adequate care, it is possible to measure protein stoichiometry (multimeric, dimeric or oligomeric) [43,46], count protein numbers [30,42,44,47,48] and characterize the spatial nano-organization of proteins [33,49]. A number of open-source (typically ImageJ-based) software is now available for biologists who wish to quantify their super-resolution images [24,37]. As the field moves forward, researchers will greatly benefit from the development of new probes or identification of buffer and imaging conditions that lead to high photon outputs, high photoactivation efficiencies and low blinking or re-activation rates.
Figure Captions

Figure 1: Dark time analysis and spatiotemporal grouping of localizations: (a) PALM image of a HeLa cell expressing SrcN15-mEos2 as a negative clustering control. (b) Plot of localizations color-coded in time. Mono-chromatic clusters are indicative of artificial clustering due to the reappearance of the same mEos2 molecule multiple times. (c) Kymograph of the spatial clusters also confirms that the localizations are clustered in time. (d-e) Setting $\tau_{\text{cutoff}} = 10$ seconds and grouping localizations that appear in time intervals shorter than $\tau_{\text{cutoff}}$ removes artificial clustering. (f) PALM image of a HeLa cell expressing $\beta_2$-AR-mEos2 as a positive control of clustering. (g-h) Spatial clusters remain even after setting $\tau_{\text{cutoff}} = 10$ seconds as expected. Reproduced with permission from reference 27.

Figure 2: Pair correlation analysis to identify spatial organization of proteins: (a) Calibration sample consisting of purified fluorescent protein randomly immobilized on glass slide. (b) PALM image of the calibration sample (PA-GFP). (c) Pair-correlation analysis to determine and correct for the contribution from multiple appearances of the same fluorescent protein (blue curve). (d-f) Application of pair correlation analysis to the PALM image of membrane proteins (red curve is before correction, blue curve is the contribution from multiple appearances of the same fluorescent protein and green curve is after correction for this effect). Reproduced with permission from reference 33.

Figure 3: Determination of fluorescent protein photoactivation (PA) efficiency using calibration samples: Drawing of the human glycine receptor (GlyR) with fluorescent protein tagged $\beta$- and untagged $\alpha$-subunit expressed in Xenopus oocyte membrane is shown. The green fluorescent protein refers to the native form of a photoconvertible fluorescent protein (e.g. mEos2) and the red fluorescent protein refers to the photoconverted form. Single step-photobleaching intensity-time traces can be obtained after photoconversion of the fluorescent proteins. Alternatively, the fluorescent proteins can be imaged with PALM to generate intensity-time traces containing peaks. Steps or peaks in intensity-time traces are counted; sufficient statistics are built and fit to the binomial distribution to obtain the photoactivation efficiency. Shown is the photoactivation efficiency of mEos2 (red curve) as a function of illumination time by 405 nm laser light to induce photoactivation. For more details, see reference 30.
References


Ref 31 uses the human Glycine receptor as a nanotemplate of known stoichiometry to characterize the percentage of successful photoactivation (photoactivation efficiency) of eight irreversibly photoactivatable and photoconvertible fluorescent proteins used in single molecule counting.


Ref 33 and 34 develop a statistical analysis method based on pair-correlation analysis to account for the effects of multiple reappearance of the same “irreversible switching” probe in super-resolution images and to determine the underlying molecular distribution and stoichiometry.


Ref 35 extends the concepts of pair-correlation analysis to reversibly switching probes and immuno-electron microscopy.


Ref 42 develops a theoretical model to determine if all the probes have been exhaustively imaged at the end of a super-resolution imaging session and uses this concept to count Gag proteins during HIV assembly.


Ref 44 uses tandem repeats of mEos2 as a calibration sample to account for the effects of overcounting (multiple reappearance of the same probe) and undercounting (failed photoactivation) to measure the changes to protein numbers on endosomes during endosome maturation.


Figure