Dissecting Oligomeric States with Photoactivated Localization Microscopy: A Numerical Model

Brian Daniels,1 Christian Wunder,2 Vanessa Chen,3 Malte Renz3*  

Abstract

Although photoactivated localization microscopy offers the potential to interrogate protein interactions in the physiological environment of a cell, uncertainties in the detection efficiency of photoactivatable proteins lead to complications with data interpretation. Here, we present a numerical model that provides probabilities to detect neighboring molecules dependent on their oligomerization status, density, detection efficiency, and radius, and can be used to assess oligomeric states or detection efficiencies of two molecular species. © 2020 The Authors.

Key terms

photoactivated localization microscopy; oligomeric states; detection efficiency; numerical model

The proper physiological function of proteins often requires assembly into multimeric complexes. Hence, it has been of major interest to gain insight into protein assembly and complex formation on the molecular level, ideally under physiologic conditions inside the living cell. Single-molecule super-resolution microscopy techniques such as photoactivated localization microscopy (PALM) (1) harbor the potential of providing quantitative information on protein assembly, including information on the number of molecules comprised in a protein complex. PALM employs genetically encoded photoactivatable fluorescent proteins (PA-FPs), which ensures a one-to-one labeling ratio of the protein of interest and the fluorescent tag. During PALM imaging, PA-FPs are stochastically and irreversibly photoactivated to a fluorescent state and subsequently irreversibly photobleached (2, 3), allowing for discrete localization and counting of detectable fluorophores.

However, such single-molecule counting techniques face challenges that complicate the faithful resolution of oligomers in any experimental sample. Two aspects of PALM particularly important to consider in this context are the blinking behavior of fluorescent proteins and the efficiency of photoactivation. Blinking of PA-FPs occurs on the timescale of milliseconds to seconds and needs to be accounted for in order to identify discrete molecular events and differentiate between the repetitive appearance of one and the same molecule versus the detection of unique molecules in the same diffraction-limited spot (1, 4). Photoactivation efficiency describes the portion of total PA-FPs that can be photoactivated to a fluorescent state. Varying photoactivation efficiencies have been described in the literature (5-8) and appear to depend on both the photophysical properties of the fluorescent protein as well as the mode of photoactivation (9). Because successful photoactivation is a prerequisite for fluorophore detection, photoactivation efficiency is a major determinant of the overall detection efficiency in single-molecule PALM imaging and remains one of the central technical challenges, alongside signal-to-noise ratio, and fluorophore quenching due to specimen preparation and fixation (5, 9-11).
Despite these challenges, different approaches to extract single-molecule numbers from PALM data have been reported. The blinking behavior of mEos2 and Dendra2 has been described and a kinetic model used to quantify single-molecule counts (12). The number of mEos2 blinking events obeys a negative binomial distribution, which can be used to assess the number of blink events for monomers and dimers (13). In another approach, the detection of discrete molecular events after accounting for blinking has been fit to binomial distributions of dimers and trimers (5). Furthermore, in larger clusters, molecule numbers and information about the oligomerization status have been derived from the amplitude of the spatial autocorrelation function (14).

Here, we provide a numerical model able to represent homogenous populations of monomers, dimers, trimers, tetramers, and pentamers of one or two different molecular species based on the probabilistic distribution of detectable fluorophores. For a given multimeric state, our model determines the probability distribution of how many detectable fluorophores (“nearest neighbors”) are expected to be encountered within a specified proximity of any given detected fluorophore, as a function only of (1) the detection radius within which a fluorophore is considered a “nearest neighbor,” (2) the detected molecular density of the experimental specimen, and (3) the overall detection efficiency of the fluorophores. The model shows distinct nearest-neighbor probability distributions for monomeric, dimeric, trimeric, tetramer, and pentameric complexes at high detection efficiency, low molecular density, and a small detection radius. These nearest-neighbor probability distributions, however, increasingly converge with decreasing detection efficiency, increasing molecular density and detection radius eventually rendering the differentiation of oligomeric complexes from random distributions of monomers impossible. We provide a parameter hierarchy showing which parameter has the greatest impact on the ability to differentiate multimeric states and determine thresholds for differentiating monomers, dimers, and trimers as a function of molecular density, detection efficiency, and detection radius. Thereby, the numerical model provides probabilities of dissecting monomers, dimers, trimers, tetramers, and pentamers for any given experimental set-up and clarifies limits of differentiating between distinct oligomeric species. Our model was also used to assess the absolute detection efficiency of several commonly used fluorescent tags with our experimental set-up and offers a method to similarly test other tags and/or other experimental set-ups.

**Materials and Methods**

**Mathematical Modeling**

Given a spatially random 2D distribution of oligomers, each of which contains N fluorophores that in turn have a probability \( P \) of being detected, our model numerically predicts the overall spatial distribution of detectable fluorophores. Clearly, the major contributors to this overall distribution are (1) the density of oligomers and (2) the degree of oligomerization, \( N \). For convention, we will refer to oligomers as “O,” detectable fluorophores as “G,” and undetectable fluorophores as “g.” Our model assumes oligomers occupy single points in space, and do not interact (i.e., do not exhibit steric hindrance).

For illustrative purposes, we first consider the effect of oligomer density on “nearest neighbor” fluorophore distributions. For a random distribution of oligomers at bulk surface concentration \( [O] \) (per unit area), the concentration of detectable fluorophores is simply given by

\[
[G] = Np[O]
\]

(1)

Provided the location of any given oligomer, the probability distribution of finding \( k \) unique oligomers within a prescribed radius, \( r \), around the oligomer of interest will obey a Poissonian distribution according to:

\[
P_G(k) = e^{-\lambda_0} \frac{\lambda_0^k}{k!}
\]

(2)

where \( \lambda_0 \) is the average number of oligomers expected to be present within the circle based on the bulk density, according to:

\[
\lambda_0 = [O] \pi r^2 = \frac{[G]}{Np} \pi r^2
\]

(3)

Monomers (\( N = 1 \)) may exist as two distinct “species”—G and g—that each have bulk concentrations of \( [G] = p[O] \) and \( [g] = (1 - p)[O] \), respectively. It, therefore, follows that the expected numbers of each of these species of oligomer within a given circle of radius, \( r \), are

\[
\lambda_G = pk_0
\]

(4)

\[
\lambda_g = (1 - p)\lambda_0
\]

(5)

Thus, the “nearest neighbor” probability distributions for each of these species will similarly be:

\[
P_G(k) = e^{-\lambda_0} \frac{\lambda_0^k}{k!} = e^{-pk_0} \frac{[p\lambda_0]^k}{k!}
\]

(6)

\[
P_g(k) = e^{-\lambda_0} \frac{\lambda_0^k}{k!} = e^{-(1 - p)\lambda_0} \frac{[(1 - p)\lambda_0]^k}{k!}
\]

(7)

For this monomeric case, the “nearest neighbor” probability distribution for detectable fluorophores is simply given by \( P_G(k) \). This result, however, is unhelpful on its own for the determination of detection efficiency, \( p \), because the overall concentration of oligomers within the sample, \( \lambda_0 \), remains unknown. We see that substituting for \( \lambda_0 \) in Eq. (6) eliminates the dependence on \( p \) as follows:

\[
P_G(k) = e^{-\lambda_0} \frac{\lambda_0^k}{k!} = e^{-[G] \pi r^2} \frac{[G] \pi r^2^k}{k!}
\]

(8)
Dissecting Oligomeric States with PALM: A Numerical Model

Thus, it remains ambiguous whether we are observing low-density monomers with high detection efficiency or high-density monomers with low detection efficiency (or anything in between).

We next consider the effect of oligomerization state, $N$, on “nearest neighbor” fluorophore distribution. At low oligomer density, $\lambda_0$ (the average number of oligomers within a radius, $r$) approaches zero, causing the probability distribution of “nearest neighbor” fluorophores to depend only on the presence of tandem fluorophores, which is a function of $N$ and $p$ via the probability mass function:

$$\begin{align*}
P_G(k) &= \binom{N-1}{k} p^k (1-p)^{N-k-1} \\
P_G(k) &= \frac{(N-1)!}{k!(N-k-1)!} p^k (1-p)^{N-k-1}
\end{align*}$$

where $N$ is the oligomerization state and $k$ is the number of detectable tandem nearest neighbors. We note that algebraically, $0^0 = 1$. It is further illustrative that at perfect detection efficiency ($p = 1$), this probability distribution reduces to:

$$P_G(k) = \frac{(N-1)!}{k!(N-k-1)!} 0^k$$

and as expected, the probability of detecting $N - 1$ nearest neighbors becomes unity. Reciprocally, at zero detection efficiency ($p = 0$), this probability distribution reduces to:

$$P_G(k) = \frac{(N-1)!}{k!(N-k-1)!} 0^k$$

and as expected, we find that the probability of detecting 0 nearest neighbors becomes unity.

Although this low-density analysis can be employed to estimate $p$ in in vitro situations where concentration can be easily titrated, this low-concentration assumption often cannot be satisfied in a cellular context. For higher-order oligomers ($N > 1$) at modest density, nearest-neighbor fluorophores within close proximity of a given fluorophore may be tandem fluorophores in the same oligomer, or fluorophores within other oligomers, or a combination of both. Therefore, our full model considers the contributions of both the oligomer density and the oligomerization state to the overall distribution of detectable fluorophores.

We next consider the case of a dimer ($N = 2$). In this case, oligomers may exist in three distinct states: GG, Gg, or gg. Oligomers have a $p^2$ probability of being in the GG state, a $2p(1-p)$ probability of being in the Gg state, and a $(1-p)^2$ probability of being in the gg state. It follows that for a given density of oligomers, the probability density of each state will distribute according to

$$P_{GG}(k) = e^{-\lambda_0} \frac{\lambda_0^k}{k!}$$

Thus, it remains ambiguous whether we are observing low-density monomers with high detection efficiency or high-density monomers with low detection efficiency (or anything in between).

We next consider the effect of oligomerization state, $N$, on “nearest neighbor” fluorophore distribution. At low oligomer density, $\lambda_0$ (the average number of oligomers within a radius, $r$) approaches zero, causing the probability distribution of “nearest neighbor” fluorophores to depend only on the presence of tandem fluorophores, which is a function of $N$ and $p$ via the probability mass function:

$$\begin{align*}
P_{GG}(k) &= e^{-2p(1-p)\lambda_0} \frac{2p(1-p)\lambda_0^k}{k!} \\
P_{Gg}(k) &= e^{-(1-p)\lambda_0} \frac{(1-p)^2\lambda_0^k}{k!}
\end{align*}$$

With the statistical distributions of each species in hand, the next step is to consider all possible circumstances that might explain the observation of $k$ nearest neighbors. For example, observing 0 fluorophores within a prescribed radius around any specific observed fluorophore ($k = 0$) can only occur when there are 0 GG or Gg oligomers present within the radius and the tandem fluorophore is also undetectable. This probability is given by the product of the probability of each of these independent events as follows:

$$P_{\text{obs}}(0) = P_{GG}(0) \times P_{Gg}(0) \times (1-p)$$

$$P_{\text{obs}}(0) = e^{-\lambda_0} \frac{\lambda_0^0}{0!} \times e^{-(1-p)\lambda_0} \frac{\lambda_0^0}{0!} \times (1-p)$$

$$P_{\text{obs}}(0) = e^{\lambda_0 - 2p\lambda_0} (1-p)$$

However, observing one fluorophore within a prescribed radius around a given fluorophore can occur when there are either: 0 GG, 0 Gg, and the tandem fluorophore is detectable; or when there are 0 GG, 1 Gg, and the tandem fluorophore is not detectable. The probability of these mutually exclusive events is given by:

$$P_{\text{obs}}(1) = P_{GG}(0) \times P_{Gg}(1) \times (1-p) + P_{GG}(0) \times P_{Gg}(0) \times (p)$$

$$P_{\text{obs}}(1) = (1-p) \times e^{-\lambda_0} \frac{\lambda_0^0}{0!} \times e^{-2p(1-p)\lambda_0} 2p(1-p)\lambda_0$$

$$+ p \times e^{-(1-p)\lambda_0} \frac{\lambda_0^0}{0!} \times e^{-2p(1-p)\lambda_0} \frac{2p(1-p)\lambda_0^0}{0!}$$

$$P_{\text{obs}}(1) = [2p\lambda_0(1-p)^2 + p] e^{\lambda_0 - 2p\lambda_0}$$

Likewise, the probability of observing two fluorophores is possible when there is 1 GG, 0 Gg, and the tandem fluorophore is not detectable; or, when there are 0 GG, 2 Gg, and the tandem fluorophore is not detectable; or, when there is 0 GG, 1 Gg, and the tandem fluorophore is detectable, which is given by:

$$P_{\text{obs}}(2) = P_{GG}(1) \times P_{Gg}(0) \times (1-p) + P_{GG}(0) \times P_{Gg}(2)$$

$$\times (1-p) + P_{GG}(0) \times P_{Gg}(1) \times (p)$$

Determining the probability distributions for higher-order oligomers and performing the subsequent “nearest neighbor” analyses are carried out in much the same way as above, albeit at much higher complexity with increasing $N$. Although the Poissonian model is mathematically unbounded ($k$ has no theoretical upper limit), we limited our sampling to

1158
with a detectable contribution from the heterodimer itself. Heterodimers plus the probability distribution of \( R \) fluorophores in a sample as a function of \( p \) [via Eq. (3)], the nearest-neighbor distribution curves ultimately depend on only three parameters: the visible density of fluorophores in a sample, \([G]\) (which can be measured experimentally), the nearest-neighbor radius, \( r \) (which is selected by the user), and the detection efficiency, \( p \) (which is unknown but can be determined by fitting our model to experimental data).

We also consider hetero-oligomers, which are oligomers comprised of multiple fluorophores with distinct emission spectra. Here, given the location of any detected fluorophore, we can consider the nearest-neighbor probability distribution of both equivalent and distinct fluorophores. For example, we can consider a heterodimer GR, which can exist in four distinct states: GR, Gr, gR, and gr. In this case, we assume green and red fluorophores will have independent detection efficiencies, \( p_G \) and \( p_R \), respectively.

When considering the spatial distribution of \( G \) in this case, we can also consider the nearest-neighbor probability distribution of \( R \), given a detected \( G \). Here, the distribution will depend on both the oligomer density plus the contribution from the heterodimer itself. Heterodimers with a detectable \( R \) will distribute according to equation [monomers] above

\[
P_{G\rightarrow R}(k) = e^{-\lambda_{G\rightarrow R} k} \frac{k^k}{k!} = e^{-p_{G} \lambda_{G\rightarrow R} k} \]  

(23)

In this case, we can also consider the nearest-neighbor distribution of \( R \), given a detected \( G \). Here, the distribution will depend on both the oligomer density plus the contribution from the heterodimer itself. Heterodimers with a detectable \( R \) will distribute according to:

\[
P_{G\rightarrow R}(k) = e^{-\lambda_{G\rightarrow R} k} \frac{k^k}{k!} = e^{-p_{R} \lambda_{G\rightarrow R} k} \]  

(24)

Therefore, finding \( k \) nearest neighbors occurs either when there are \( k \) \( R \)-detectable oligomers AND the tandem fluor is undetectable, or when there are \( k - 1 \) detectable oligomers AND the tandem fluor is detectable as follows:

\[
P_{\text{obs},G}(1) = P_{G\rightarrow R}(k) \times (1 - p_R) + P_{G\rightarrow R}(k-1) \times p_R \]  

(25)

At low heterodimer density, the probability of finding one nearest neighbor of the opposite color approaches the detection efficiency of that color since the contribution from other heterodimers becomes negligible, as follows:

\[
P_{\text{obs},R}(1) \big|_{\lambda_{G\rightarrow R}=0} = p_R \]  

(26)

and thus also

\[
P_{\text{obs},G}(1) \big|_{\lambda_{G\rightarrow R}=0} = p_G \]  

(27)

These special cases were used to validate the detection efficiencies predicted with our homo-oligomer model by using low oligomer concentrations in vitro.

These basic concepts can be similarly extended to higher-order hetero-oligomers such as GGR, GRR, and so on and used to determine the absolute detection efficiencies of fluorophores (which may depend on the local environment, method of linkage to neighboring molecules, the composition of neighboring molecules themselves, etc.) as well as to make inferences regarding the actual stoichiometry of fluorophores when only a portion of them are experimentally detectable.

**Bacterial Plasmids and Recombinant Protein Expression, Purification**

mGFP-PAmCherry1 and mGFP-mCherry1 fusion constructs were amplified as a *Ndel-Notl* fragment, using the N-terminal primer 5'-ATT AAC ATA TGG TGA GCA AGG GGC AGG AG containing an *Ndel* site (underlined) and the C-terminal primer 5'-ATA ATT GAA TTC GTC TAG AGT CTC GGC CG 3' containing a *Notl* site (underlined), and inserted into the multiple cloning site of the pHis-parallel-2 vector. The stop codon following the fluorescent protein chimera was deleted.

Fluorescent proteins were expressed in a Rosetta2-pLysS strain (Novagen) in LB media. An overnight pre-culture (37°C) was diluted 1:1,000 in a 4 L bacterial culture flask with six side and bottom baffles using 800 ml LB media. This bacterial culture was grown at 37°C, 300 rpm until OD600 reached 0.6 and cooled down to 25°C. From now on, every expression and purification step was done under strict protection from any light source. Fluorescent protein expression was induced with 8 μM IPTG at 25°C for 14 h. Bacterial cell pellets were resuspended in ice-cold PBS pH 7.3/10% glycerol/protease inhibitor cocktail (Roche) and sonicated on ice (Sonicator, ultrasonic processor, XL2020, Misonix). Subsequent cell lysates were centrifuged at 43,000g for 60 min at 4°C. The soluble fraction was incubated with NTA-agarose (Superflow NTA, Qiagen), washed with 300 ml PBS/10% glycerol/20 mM imidazole, eluted in PBS/500 mM imidazole and filtrated by using a 0.22 μm PVDF membrane (Millipore, Millex-GV). 95% purity by Coomassie staining was achieved using size-exclusion chromatography (HiLoad 16/600 Superdex200, without UV detection) in PBS on an Äkta protein purification FPLC system. Peak fractions were detected at 280 nm (NanoDrop). Small single-use aliquots were snap-frozen in liquid nitrogen and stored at −80°C. The maturation of fluorescence proteins was optimized beforehand based on expression temperature, IPTG concentration, and duration of IPTG induction.

**Photoactivated Localization Microscopy**

Glass slides were washed by boiling in acetone for 10 min at 75°C, then in a mixture of 1:1.5 H2O2/NH2OH/H2O for
**RESULTS**

**Numerical Model—Parameters**

In our numerical model, oligomer distributions are assumed to follow a two-dimensional Poisson distribution. The fluorophore nearest-neighbor probability distribution is calculated with the following defining independent parameters: detection efficiency (experimentally measured or based on published data), visible molecular density (experimentally measured), and radius within which a neighboring molecule is detected (specified). Numerical values for the actual molecular density in the sample and the probabilities of a neighboring molecule for a monomer, dimer, trimer, tetramer, and pentamer of one and two different molecular species will be calculated (Supplementary Material).

**Differentiating Homodimers from Monomers**

In a sparse distribution of homodimers, for example, 3 molecules/μm² and with a detection efficiency for the experimental set-up of 95%, almost every molecule of a homodimer has a neighboring molecule within a 15 nm radius while random distributions of monomers of the same molecular density do not have any neighboring molecules (Fig. 1A). The probabilities of having one neighboring molecule based on our numerical model are 2.1% for a monomer and 94.9% for a homodimer. If, however, the distribution of molecules in a sample is denser, for example, 90 molecules/μm², the detection efficiency drops to 10%, and the detection radius increases to 60 nm, the probability of having a neighboring molecule for a homodimer decreases to 35.2%. At the same time, the probability for a monomer to have a neighboring molecule just by chance increases to 36.8%. Thus, under these conditions, the probabilities of finding neighboring molecules for monomers and homodimers are basically identical and homodimers cannot be reliably separated from monomers (Fig. 1B). The same applies to the differentiation of homotrimers, homotetramers, and homopentamers (Fig. 1C,D).

**Impact of Molecular Density, Detection Efficiency and Radius on Differentiating Oligomers**

Supplemental Figure 1 illustrates how changing one parameter, that is, the detection efficiency, molecular density, or detection radius, while keeping the other parameters constant, affects the ability to differentiate monomers, dimers, and trimers. In general, the denser the molecule distribution, the lower the detection efficiency, and the larger the detection radius, the more the nearest-neighbor distribution curves for monomers, dimers, and trimers converge. To quantify which parameter has the greatest impact on differentiating oligomeric states, we calculated the difference in nearest-neighbor probabilities for monomers and dimers per percent change of each parameter over a specific range; for the detection efficiency, a range of 1–100% was used, for the molecular distribution, a density of 1–100 molecules/μm², and for the detection radius, a range from 10 to 60 nm. While a parameter was changed stepwise by 1%-increments, the other parameters were kept constant at 50% detection efficiency, 50 molecules/μm², and a radius of 35 nm, respectively. Resulting graphs are shown in Figure 1E. An increase in detection efficiency, decrease in molecular density, and detection radius result in a distinct nonlinear increase in the ability to differentiate monomers and dimers. Over the prespecified range, the percent change in detection efficiency has the greatest impact, and a change in the detection radius the smallest. In contrast to detection efficiency and molecular density, the impact of the detection radius diminishes per percent change with decreasing radii and the resulting graph displays an asymptotic behavior (Supplemental Figure 2).

**Thresholds for Differentiating Monomers, Dimers, and Trimers**

To provide thresholds for a reliable separation of oligomers and monomers, we calculated the difference in probability for one nearest neighbor for monomers and homodimers and the difference in probability for two nearest neighbors for homodimers and homotrimers as a function of detection efficiency (x-axis) and molecular density (y-axis). Resulting surface plots are shown in Figure 1F and Supplemental Figure 3. To reduce the parameter space to a three-dimensional space, the detection radius was set to 30 nm and 40 nm, respectively. The choice of these radii is based on the published uncertainty of the molecular localization, σ, ranging from 10 to 25 nm (1, 11, 15, 16). If the detection radius were to be...
reduced below the uncertainty of the molecular localization, the co-localization of oligomeric proteins would be missed. For a reliable differentiation of monomers, dimers, and trimers, a 10%-difference in the probability of having one or two neighboring molecules, respectively, was here considered sufficient. As shown in Figure 1F, to differentiate monomers
and dimers with, for example, a 10% detection efficiency would only be possible up to a molecular density of 1 molecule/μm². To be able to differentiate monomers and dimers at a molecular density of 100 molecules/μm² requires at the minimum a 23% detection efficiency for a 30 nm detection radius and 42% detection efficiency for a 40 nm detection radius. To also be able to differentiate dimers from trimers and not only monomeric distributions from some sort of lower-order oligomers, a 38% detection efficiency at 100 molecules/μm² for a 30 nm radius and a 45% detection efficiency for a 40 nm radius. Thus, to not only reliably differentiate monomers from dimers, but also dimers from trimers the minimum detection efficiency required to separate these oligomers is 38% and 45% for a molecular density of 100 molecules/μm². Thus, to not only reliably differentiate monomers from dimers, but also dimers from trimers the minimum detection efficiency required to separate these oligomers is 38% and 45% for a molecular density of 100 molecules/μm² at a radius of 30 and 40 nm, respectively.

Differentiating Hetero-oligomers
For dimeric or trimeric complexes of two different molecular species labeled with two different colors, it is important to consider that photoactivation and detection efficiency of different photoactivatable fluorescent proteins can be different. Therefore, the detection efficiency of different colors can be independently changed in our numerical model. Figure 2A,B shows a scenario where the detection efficiency of one color is 95% while the detection efficiency of the other color is only 50% in a sparse molecular distribution of 3 molecules/μm². Calculated nearest-neighbor probability curves for monomers, heterodimers, and heterotrimers are displayed.

Detection Efficiencies of the Fluorescent Proteins GFP, Cherry, and PA-GFP
So far, we have used the numerical model to determine the oligomerization status of fluorescently labeled proteins with measured molecular density and specified detection efficiency and radius. If, however, the detection efficiency of the fluorescent protein in the specific experimental set-up is unknown, fluorescent proteins of known oligomerization status can be used to assess the detection efficiency of a fluorescent protein with the numerical model. We show examples of purified fluorophore chimeras that are known heterodimers, that is, the GFP—Cherry and PA-GFP—Cherry fusion proteins. The fluorescent proteins of these chimeras will have different detection efficiencies. The purified chimeras were sparsely dispersed on a glass surface in a molecular density of less than 1 molecule/μm². Then, they were imaged and localized to test for the detection efficiency of the different fluorescent proteins. For the always-on fluorescent protein chimera GFP—Cherry, 53 green molecules, 42 red molecules were detected. 35 molecules had a spectrally distinct neighbor. Figure 2C,D depicts the experimental data and theoretical probabilities of neighboring molecules calculated with the numerical model. The experimental data were best matched with a detection efficiency of 69.8% for red molecules and 84.1% for green molecules. Mean detection efficiencies were 71.1 ± 5.9% and 86.5 ± 4.3%, respectively (n = 5 each). Figure 2C,D also shows the calculated random distribution of red and photoactivatable green molecules for the same detection efficiencies, molecular density, and detection radius. The distinct detection efficiencies of the always-on fluorescent proteins with a lower detection efficiency Cherry is consistent with the less efficient maturation of red fluorophores reported in the literature (17).

For the PA-GFP—Cherry dimer, in total 60 red molecules were detected in one field of view and subsequently, 15 PA-GFP molecules were switched on to fluoresce so that 15 red molecules had a green neighbor while 45 had no green neighbor. Figure 2E shows the experimental and theoretical data for a detection efficiency of 25% for PA-GFP. Mean detection efficiency was 27.6 ± 0.8% (n = 15) as previously published (9). Again, the random distribution of red and photoactivatable green molecules for the same detection efficiencies, molecular density, and detection radius is shown as well. Beyond these examples, the model can be applied to coupled PA-FPs of higher molecular density.

Discussion
The here presented numerical model provides a tool to dissect monomers, dimers, trimers, tetramers, and pentamers based on the primary data, that is, the discrete localization of single molecular events, obtained with PALM. It provides a numerical basis for the correct interpretation of primary PALM data and helps benefit from one of the main advantages of PALM over other single-molecule techniques such as step-wise photobleaching.
While PALM can in principle be used to assess higher-order oligomers (tetramers, pentamers, etc.) in densely expressing cells, step-wise photobleaching is limited to the resolution of dimers and trimers in low-expressing cells.

The presented model helps clarify relevant parameters for resolving monomers through pentamers, that is, detection efficiency, molecular density and detection radius, and how these parameters are interrelated. Over a specified range,
detection efficiency has the highest impact on differentiating oligomers and detection radius the lowest impact.

Many factors contribute to the detection efficiency of any experimental set-up: the signal-to-noise ratio, sample preparation (resulting in fluorophore quenching), exhaustive recording of single molecules, etc. However, even under ideal experimental conditions, the detection efficiency will be limited by properties of the fluorescent proteins such as protein folding and maturation. Inefficiencies in protein folding and maturation have been well described in the literature and are illustrated here by distinct detection efficiencies of the always-on fluorescent proteins GFP and Cherry (17). A PA-FP specific feature, that is, photoactivation, further contributes to detection efficiency. And while a rather broad range of photoactivation efficiency has been published for various PA-FPs, it is agreed upon that photoactivation efficiency is not 100%. In addition, photoactivation efficiency may in fact depend upon the illumination mode (9). With our model, minimal detection efficiencies required for the determination of oligomeric states can be defined as a function of molecular density and detection radius. To resolve, for example, with a 10% difference in probability monomers from dimers, and dimers from trimers at a density of 100 molecules/μm² with a 30-nm detection radius, the minimal detection efficiency required is 38%.

Regarding the molecular density, it is important to realize that some molecules are expressed in a cell but not detected. Thus, a visible and an actual molecular density need to be taken into account, even if every molecule has been labeled with a PA-FP and no endogenous unlabeled pool exists in the biological sample. The higher the actual molecular density, the higher the chance of random spatial overlap of noninteracting monomeric molecules.

The detection radius is another independent variable of our model and can be freely set. It may be considered that for increasingly smaller radii the uncertainty of localization, σ, obtained in a PALM experiment may confound the results. The uncertainty of localization σ has been defined as the standard deviation of multiple localization measurements of a single object (1, 18). Due to basic descriptive statistics, the mathematically determined localization of a molecule plus/minus the localization uncertainty multiplied by 1.96 equals 95% of the possible localization of the actual single object. Therefore, when using the mathematical model presented here the detection radius may not be set too small in order to not miss neighboring molecules, which are part of an oligomeric complex.

In summary, the provided numerical model will help prevent misinterpretation of primary PALM data, provides thresholds for detection efficiency required to resolve monomers from oligomers, and permits the differentiation of monomers, dimers, trimers, tetramers, and pentamers in densely expressing cells.

**Literature Cited**

1. Betzig E, Patterson GH, Sougrat R, Lindwasser OW, Olenych S, Bonifacino JS, Davidson MW, Lippincott-Schwartz J, Hess HF. Imaging intracellular fluorescent proteins at nanometer resolution. Science 2006;315(5816):1642-1645.
2. Patterson GH, Lippincott-Schwartz J. A photoactivatable GFP for selective photolabeling of proteins and cells. Science 2002;297(5588):1873-1877.
3. Subach FV, Patterson GH, Manley S, Gillette JM, Lippincott-Schwartz J, Verkhusha VV. Photoactivatable mCherry for high-resolution two-color fluorescence microscopy. Nat Methods 2009;6(2):153-159.
4. Annibale P, Vanni S, Scarselli M, Rohlsberger U, Radenovic A. Quantitative photo activated localization microscopy: Unraveling the effects of photoblinking. PLoS One 2011;6(7):e22678.
5. Durisic N, Laparra-Cuervo L, Sandoval-Alvarez A, Borbely JS, Lakadamyali M. Single-molecule evaluation of fluorescent protein photoactivation efficiency using an in vivo nanotemplate. Nat Methods 2014;11(2):156-162.
6. Wiedenmann J, Iwaienke S, Oswald F, Schmitt F, Rocker C, Saliå A, Spindler KD, Nirenhaus GU. EosFP, a fluorescent marker protein with UV-inducible green-to-red fluorescence conversion. Proc Natl Acad Sci U S A 2004;101(45); 15905-15910.
7. Habuchi S, Tsurus H, Kochaniak AB, Miyawaki A, van Oijen AM. mKikGR, a monomeric photoswitchable fluorescent protein. PLoS One 2008;3(12):e3944.
8. McKinney SA, Murphy CS, Hazelwood KL, Davidson MW, Looger LL. A bright and photostable photoconvertible fluorescent protein. Nat Methods 2009;6(2):131-133.
9. Renz M, Wunder C. Internal rulers to assess fluorescent protein photoactivation efficiency. Cytometry Part A 2018;93(4):411-419.
10. Renz M, Lippincott-Schwartz J. Single-molecule counting with PALM. Biophys J 2011;100(3):350.
11. Renz M, Danieri BR, Vamosi G, Arias IM, Lippincott-Schwartz J. Plasticity of the asialoglycoprotein receptor deciphered by ensemble FRET imaging and single-molecule counting PALM imaging. Proc Natl Acad Sci U S A 2012;109(44):E2989-E2997.
12. Lee SH, Shin JY, Lee A, Bustamante C. Counting single photoactivatable fluorescent molecules by photoactivated localization microscopy (PALM). Proc Natl Acad Sci U S A 2012;109(43):17436-17441.
13. Fricke F, Beaudouin J, Eils R, Heilemann M. One, two or three? Probing the stoichiometry of membrane proteins by single-molecule localization microscopy. Sci Rep 2015;5:14072.
14. Sengupta P, Jovanovic-Talisman T, Skoko D, Renz M, Veatch SL, Lippincott-Schwartz J. Probing protein heterogeneity in the plasma membrane using PALM and pair correlation analysis. Nat Methods 2011;8(11):969-975.
15. Shroff H, Galbraith CG, Galbraith JA, White H, Gillette J, Olenych S, Davidson MW, Betzig E. Dual-color superresolution imaging of genetically expressed probes within individual adhesion complexes. Proc Natl Acad Sci U S A 2007;104(51):20308-20313.
16. Manley S, Gillette JM, Patterson GH, Shroff H, Hess HF, Betzig E, Lippincott-Schwartz J. High-density mapping of single-molecule trajectories with photoactivated localization microscopy. Nat Methods 2008;5(2):155-157.
17. Balleza E, Kim JM, Chand P. Systematic characterization of maturation time of fluorescent proteins in living cells. Nat Methods 2018;15(1):47-51.
18. Huang B, Babcock H, Zhuang X. Breaking the diffraction barrier. Super-resolution imaging of cells. Cell 2010;143(7):1047-1058.