Dissertation

submitted to the

Combined Faculty of Natural Sciences and Mathematics

of the Ruperto Carola University Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

Presented by

Klaus Yserentant, M. Sc.

born in Aachen

Oral examination: 11.03.2022

Tools for quantitative Fluorescence Microscopy on the Nanoscale

Referees: Prof. Dr. Dirk-Peter Herten
Priv. Doz. Dr. Richard Wombacher

Abstract

Biology at the molecular level is driven by macromolecules interacting with each other to form multi-component complexes. To understand the function of macromolecular complexes, quantitative information about their composition and dynamics are essential. Single-molecule fluorescence microscopy offers unique possibilities to provide such information, but relies on the labeling of target molecules with fluorescent markers. In addition, the fluorescence signal emitted by a sample does not readily inform about absolute fluorophore numbers. To measure complex stoichiometries with fluorescence microscopy, it is therefore required to know the fraction of successfully labeled target molecules and to apply a method which infers the number of fluorescent markers from the recorded fluorescence signal. To approach these challenges, I have established and optimized tools to allow for a reliable determination of labeling efficiencies and secondly to successfully determine fluorophore numbers in fluorescently labeled complexes inside cells.

In the first part of this study, I show that the labeling efficiency achieved with self-labeling protein tags in living and fixed cells can be calibrated by single-molecule colocalization analysis. An improved and validated data processing pipeline enabled me to systematically study the performance of the self-labeling protein tags, $\mathrm{SNAP_t}$-tag and HaloTag across different labeling conditions and with different fluorescent ligands. I found that labeling efficiencies for both tags depend on the ligand used and are limited to sub-stoichiometric levels for all tested ligands and labeling conditions. The highest labeling efficiency for either tag was achieved by labeling of $\mathrm{SNAP_t}$-tag with BG-SiR in fixed cells, where a maximum labeling efficiency of 65 % was reached. The developed calibration approach provides a generalizable platform for the development and benchmarking of new labeling schemes for quantitative fluorescence microscopy.

To address the need for a method to translate fluorescence intensity into absolute fluorophore numbers, `quickPBSA` was established as a new framework for fluorophore counting by photobleaching step analysis. `quickPBSA` was validated with simulations and in vitro measurements on DNA origami demonstrating an accessible counting range of up to 35 fluorophores and a 100-fold improvement in computational cost compared to previous algorithms. I could show that by combining improved data acquisition conditions with the `quickPBSA` framework for photobleaching step detection enables measuring stoichiometries on complexes labeled with up to 32 fluorophores inside cells.

In combination, the developments presented in this study provide a comprehensive approach for measuring stoichiometries of protein complexes in situ and put applications in cell biology into perspective.

Zusammenfassung

Biologische Prozesse werden auf molekularer Ebene von Komplexen, welche durch die Interaktion verschiedener Makromoleküle entstehen, angetrieben. Um die Funktion dieser Komplexe zu verstehen, werden quantitative Informationen über ihre Zusammensetzung und Dynamik benötigt. Einzelmolekülfluoreszenzmikroskopie ist in der Lage diese Informationen zu erheben. Eine Herausforderung hierbei ist jedoch, dass hierfür eine fluoreszente Markierung der Zielproteine notwendig ist. Gleichzeitig gibt das Fluoreszenzsignal, das in einer Probe detektiert wird nicht direkt Auskunft über die Anzahl der enthaltenen Fluoreszenzmarker. Um die Stöchiometrie von Komplexen mittels Fluoreszenzmikroskopie zu bestimmen ist es deshalb unerlässlich den Anteil der erfolgreich markierten Zielmoleküle zu bestimmen und außerdem eine Methode anzuwenden, welche es erlaubt die Anzahl an fluoreszenten Markern in einer Probe mittels des gemessenen Fluoreszenzsignals zu bestimmen.

Um diese methodischen Herausforderungen anzugehen, habe ich im Rahmen der vorliegenden Arbeit Methoden zur Bestimmung der Markierungseffizienz, sowie zur Messung der Anzahl von Fluoreszenzmarkern mittels Einzelmolekülfluoreszenzmethoden etabliert und optimiert. Im ersten Teil dieser Arbeit habe ich gezeigt wie die Markierungseffizienz von den selbst-markierenden Protein-Tags SNAP$_\mathrm{f}$-tag und HaloTag in lebenden und fixierten Zellen mit Hilfe von Einzelmolekülkolokalisationsanalyse kalibriert werden kann. Eine optimierte und validierte Plattform zur Datenanalyse haben es mir erlaubt die Leistungsfähigkeit der beider Protein-Tags unter verschiedenen Bedingungen und für verschiedene fluoreszente Liganden systematisch zu untersuchen. Dabei konnte ich zeigen, dass die Markierungseffizienz fuer beide Protein-Tags von dem jeweilig verwendeten Liganden abhängt und für alle getesteten Bedingungen auf sub-stöchiometrische Werte limitiert ist. Die höchste Markierungseffizienz von 65 % wurde fuer die Markierung mittels SNAP$_\mathrm{f}$-tag unter Verwendung des Liganden BG - SiR und in fixierten Zellen erreicht. Der hier vorgestellte Kalibrationsansatz kann als flexible Plattform zur Charakterisierung und Entwicklung neuer

Ansätze zur Fluoreszenzmarkierung für die quantitative Fluoreszenzmikroskopie eingesetzt werden.

Im zweiten Teil der Arbeit habe ich mich auf die Weiterentwicklung von Methoden, die es erlauben aus dem gemessenen Fluoreszenzsignal die Anzahl an Fluorophoren zu bestimmen, fokussiert. `quickPBSA` wurde als ein Verfahren entwickelt, welches das Zählen von Fluoreszenzmarkern mittels Beobachtung schrittweiser Photozerstörung ermöglicht. Hierzu habe ich `quickPBSA` zunächst mit Hilfe von Simulationen, sowie in vitro Messungen auf DNA-Origami validiert und konnte so zeigen, dass bis zu 35 einzelne Fluoreszenzmarker zuverlässig identifiziert werden können. Im Vergleich zu existierenden Algorithmen benötigte `quickPBSA` zudem eine 100-fach geringere Rechenzeit. Schließlich konnte ich zeigen, dass `quickPBSA` in Kombination mit optimierten Bedingungen fuer die Datenaufnahme in Zellen, die Stöchiometrie von Komplexe mit bis zu 32 Fluoreszenzmarkern auch direkt in Zellen erfolgreich bestimmen kann. In Kombination, stellen die in dieser Arbeit beschriebenen Entwicklungen einen umfassenden Ansatz zur Bestimmung der Stöchiometrie von Proteinkomplexen in situ dar und eröffnen damit die Möglichkeit für zukünftige Anwendungen in der Zellbiologie.

Für Julia.

Acknowledgments

The work presented in this study would not have been possible without the guidance and support of my colleagues, friends and family. I would like to take this opportunity to thank all of you for having supported me during the past years.

Dirk, I would like to begin by thanking you for providing me the opportunity and the freedom to work on my ideas and for supporting me throughout the time in your lab. Thank you for always providing words of encouragement, guidance and for keeping me motivated to get things done.

Richard, thank you for having become a close collaborator over the past years and for always providing feedback, ideas and a valuable chemical biology perspective.

Felix and Johan, thank you for having worked with me on the developments shown in this thesis. Your expertise and ideas helped shape the projects we worked on and it has been fantastic working with you. Beyond this, thank you for having been the best office mates imaginable. For coffee breaks, discussions, Sommerfeste and conference trips to Lausanne, Bruges, Berlin, Rennes, and Oxford.

Philipp, thank you for coffee breaks during the past two years. For your positive view on the world, for providing critical support during the long-distance submission of this thesis, for paper submission drinks, and of course, for two phenomenal collaborative projects.

Kristin, thank you for introducing me to the field of molecular counting, for great discussions and for having worked with me and for continued collaborations even after your move to Lausanne.

Sigi, thank you for your work on the calibration probe project, for invaluable coffee expertise and for beers on bridges.

Michi, thank you for a fantastic Mandi collaboration and for putting up with me during our paper revision marathon.

Jonas, thank you for joining me in working on our CLEM project, for cakes and your help with last-minute measurements.

Thank you to all former and current members of the Herten lab. Felix, Johan, Kristin, Sigi, Jonas, Dodo, Andreas, Barzi, Anne, Wioleta, Tanja, Yin and Stan for having made the time in our lab an unforgettable experience.

Further, I want to thank

... my former rotation students Michael, Lukas, Jakob, Leonard, Kaisa, Antos, Christoph, Philipp and Theresa for their work and enthusiasm.

... the current and former members of the Schröder lab: Rasmus, Irene, Chris, Ira, Götz, Johan, Setsuko, Anne and Katharina for support during the initial phase of my PhD.

... my further external collaborators, Fabian Erdel, Valentin Dunsing, Vytautas Navikas, Volker Middel, Miao Yu, Tina Wiegand and Georg Sedelmeier for great collaborations.

... Ursula Kummer and Petr Chlanda for joining my thesis evaluation committee.

Last but not least, thank you to my family and friends, especially my parents Inge and Harry, my brother Max and my friends Wiebke, Moritz, Jakob, Philipp, Dominik and Joseph. I owe my biggest gratitude to all of you for supporting me on so many different levels. Thank you for listening to me, for keeping me going, for being patient while I finished the final parts of this thesis and for sharing so many wonderful moments.

Finally, Julia. Thank you for being who you are, for always being by my side, for being my best friend and fiercest critic. Thank you for having been part of so many great experiences in Heidelberg and around the world. Thank you for supporting me since day one and particularly in the final weeks of writing this thesis. Thank you... for so much more.

Peer-reviewed Publications

* denotes equal contribution as first authors.

12. **Yserentant K***, Hild F*, Hänselmann S*, Chmielewicz W, Hummert J, Salopiata F, Fackler O, Klingmüller U, Herten DP. A general method for measuring the absolute degree of labeling for in-situ experiments with self-labeling protein tags. *Manuscript in preparation*.

11. Hild F*, Werther P*, **Yserentant K**, Wombacher R, Herten DP. A dark intermediates in the fluorogenic reaction between tetrazine fluorophores and trans-cyclooctene. *Manuscript in preparation*

10. Ziegler MJ*, **Yserentant K***, Dunsing V, Middel V, Gralak AJ, Pakari K, Bargsted J, Kern C, Petrich A, Chiantia S, Strähle U, Herten DP, Wombacher R. Mandipropamid as a chemical inducer of proximity for in vivo applications. *Nature Chemical Biology* **18**, 64-69 (2022).

9. Hummert J*, **Yserentant K***, Fink T, Euchner J, YX Ho, SA Tashev, Herten DP. Improved photobleaching step analysis for robust determination of protein complex stoichiometries. *Molecular Biology of the Cell* **32**, 21 (2021).

8. Werther P*, **Yserentant K***, Braun F, Grußmayer KS, Navikas V, Yu M, Zhang Z, Ziegler MJ, Mayer C, Gralak AJ, Busch M, Chi W, Rominger F, Radenovic A, Liu X, Lemke EA, Buckup T, Herten DP, Wombacher R. Bioorthogonal red and far-red fluorogenic probes for wash-free live-cell and super-resolution microscopy. *ACS Central Science* **7**, 1561-1571 (2021).

7. Navikas V*, Leitao SM*, Grußmayer KS, Descloux A, Drake B, **Yserentant K**, Werther P, Herten DP, Wombacher R, Radenovic A, Fantner GE. Correlative 3D

microscopy of single cells using super-resolution and scanning ion-conductance microscopy. *Nature Communications* **12**, 4565 (2021).

6. Sedlmeier G, Al-Rawi V, Buchert J, **Yserentant K**, Rothley M, Hurrle T, Gräßle S, Steshina A, Decraene C, Wu R, Thiele W, Utikal J, Garvalov BK, Abuillan W, Tanaka M, Herten DP, Hill C, Jung N, Bräse S, Sleeman JP. Id1 and Id3 are regulated through matrix-assisted autocrine BMP signalling and represent therapeutic targets in melanoma. *Advanced Therapeutics* **4**, 2000065 (2021).

5. Erdel F, Rademacher A, Vlijm R, Tünnermann J, Frank L, Weinmann R, Schweigert E, **Yserentant K**, Hummert J, Normand C, Bauer C, Schumacher S, Al Alwash A, Herten DP, Engelhardt J, Rippe K. Mouse heterochromatin adopts digital compaction states without showing hallmarks of HP1-driven liquid-liquid phase separation. *Molecular Cell* **78**, 236–249.e7 (2020).

4. Werther P*, **Yserentant K***, Braun F, Kaltwasser N, Popp C, Baalmann M, Herten DP, Wombacher R. Live-Cell Localization Microscopy with a Fluorogenic and Self-Blinking Tetrazine Probe. *Angewandte Chemie* **59**, 804-810 (2020).

3. Wiegand T, Fratini M, Frey F, **Yserentant K**, Liu Y, Weber E, Galior K, Ohmes J, Braun F, Herten DP, Boulant S, Schwarz US, Salaita K, Cavalcanti-Adam AE, Spatz JP. Forces during cellular uptake of viruses and nanoparticles at the ventral side. *Nature Communications* **11**, 32 (2020).

2. Grußmayer KS*, **Yserentant K***, Herten DP. Photons in – numbers out: perspectives in quantitative fluorescence microscopy for in situ protein counting. *Methods and Applications in Fluorescence* **7**, 012003 (2019).

1. Schelkle KM*, Schmid, C*, **Yserentant K**, Bender M, Wacker I, Petzoldt M, Hamburger M, Herten, DP Wombacher R, Schröder RR, Bunz UHF. Cell fixation by light-triggered release of glutaraldehyde. *Angewandte Chemie* **129**, 4802-4806 (2017).

Contents

Part I

Introduction

Chapter 1

Motivation

Biological processes are driven by billions and billions of individual macromolecules co-existing and interacting to form complex multicomponent structures. Among this pool of macromolecules are many different molecular species belonging to different classes of macromolecules including nucleic acids, proteins and lipids. Higher order structures such as subcellular organelles and cells made up of individual macromolecules then interact to give rise to life on the macroscopic scale[1,2]. Biomedical research seeks to understand this intricate network of different types of molecules by mapping their molecular properties, their interactions and their distribution across space and time.

One organizing principle in biology is that macromolecules can form complexes by self-interaction to form homotypic complexes or by interaction with other molecular species to form heterotypic complexes. These complexes vary with respect to their stability, the number of individual components and copy numbers per component[3,4]. One example for a stable macromolecular complex is the Nuclear pore complex (NPC), a multi-protein machinery made up of $\sim 1 \times 10^3$ protein molecules from at least >30 different protein species, which form a channel with selective permeability in the nuclear membrane of eukaryotic cells[5,6]. NPCs reach a size of $>100\,\mathrm{nm}$ in diameter and a molecular weight of $>1 \times 10^2$ MDa and were recently shown to exhibit a higher degree of structural and functional flexibility than previously known[7,8].

Transmembrane receptor clusters in the plasma membrane of cells are another prominent class of macromolecular complexes. Receptor clusters are also often composed of several different molecular species and multiple copies of each molecular species, but tend to exhibit a more dynamic behavior. The specific function of receptor clusters can vary, but in general, they form to enable, regulate or modulate signal transduction across the plasma membrane[9–11]. For example, the surface receptors of immune cells are often organized into large multicomponent assemblies, which can change in terms of size and composition upon contact between the immune cell and its cognate antigen. This compositional flexibility enables the highly sensitive, yet specific response that is the basis of the immune system in many multicellular organisms[12–15].

Despite decades of intense research, many questions regarding the specific composition, formation and dynamic regulation of macromolecular complexes in cells remain open[16,17]. Since macromolecular complexes are too small to be directly observed by eye, analytical techniques with sufficient sensitivity and selectivity to study their composition by indirect means are required. In that respect, fluorescence microscopy offers unique possibilities as it allows to study individual species of biomolecules with molecular specificity even in the presence of high concentrations of structurally similar molecules. The extraordinary sensitivity of modern fluorescence microscopes enables observing individual molecules in real time and with high spatial resolution.

As most biological macromolecules are not fluorescent, fluorescence microscopy typically requires conjugation of a fluorescent label to the macromolecule of interest. For this purpose, different types of fluorescent labels with vastly different properties have been developed[18]. As a consequence, fluorescent labeling is the basis for the extraordinary specificity of fluorescence microscopy. However, this also means that if the fluorescence signal emitted from labels is used as proxy for the macromolecule under investigation, the fraction of target molecules carrying a fluorescent label has to be calibrated. Due to the diffraction of light, our ability to directly count the number of fluorescent labels contained in a macromolecular complex by enumeration of fluorescent signals in images is limited. In order to obtain information about the composition and stoichiometry of macromolecular clusters, it is therefore necessary to use additional information contained in the fluorescence signal[19].

In the course of the work presented in this study, I will address both, the need for calibrated labeling of target molecules and the need for a method to convert fluorescence signals into label numbers. This work was done with the goal to provide a workflow which enables researchers to elucidate the composition of molecular complexes with single-molecule sensitivity and in their native cellular context.

Chapter 2

Background

Fluorescence microscopy has become an essential technique for biological research. In order to maximize the information obtained from fluorescence microscopy experiments, it is important to be aware of the basic principles and limitations of fluorescence and image formation in a fluorescence microscope. In the next chapter, I will provide a brief introduction to fluorescence and fluorescence microscopy. I will then briefly discuss how POIs can be labeled using different fluorescent markers and how emitter number estimates can be obtained from fluorescence microscopy. Finally, I will provide a brief overview of reference standards that have been used for calibration and validation of emitter counting fluorescence microscopy. For a detailed description of the phenomenon of fluorescence and its applications as analytical technique, the interested is referred to books covering both topics in great detail[20,21]

2.1 Basics of Fluorescence Microscopy

Fluorescence is defined as the emission of photons by molecules upon return from an electronically excited state into their ground state. Typically, this transition occurs by relaxation from the first excited state S_1 to the ground state S_0. Excitation of molecules into a higher energy electronic state occurs through absorption of electromagnetic radiation. In the case of fluorescence microscopy and spectroscopy, this is typically achieved by means of shining visible light onto the sample. In a seminal discovery in 1852, Stokes observed that some substances emit light with a red-shifted spectrum compared to the incident light[22]. It was later discovered that this effect is due to the dissipation of energy in the excited electronic state during return to the lowest vibrational state. Here, the transition between vibrational states occurs on time scales shorter than the time scale for S_1-S_0 transition by emission of a photon. This is the reason why the wavelength of light emitted as fluorescence is independent from the wavelength of light absorbed during excitation (Kasha's rule)[23].

In addition to the S_1 - S_0 transition via emission of a photon, several non-radiative transition mechanisms such as internal conversion between vibrational states of different electronic states can occur. Intersystem crossing from S_1 to the triplet excited state T_1 requires spin conversion and therefore typically occurs with lower probabilities. Due to the low probability of the T_1 - S_0 transition the residence time in the triplet state is typically substantially longer than in S_1. The sequestration of a fluorescent molecule in T_1 away from the fast cycling between S_0 and S_1 results in a temporary loss in fluorescence signal from this molecule and is also known as triplet blinking.

Normally, non-radiative return to the ground state S_0 is favored in solution, but radiative return to S_0 in a process termed phosphorescence can occur in rigid media.

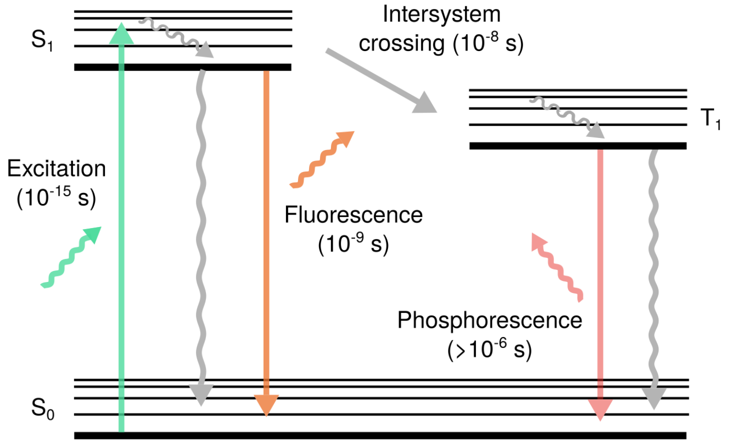

Figure 2.1 Simplified Jablonski diagram to illustrate state transitions of fluorescent molecules. Transitions between electronic ground state (S_0), first excited state (S_1) and triplet state (T_1) and typical time scales for individual processes. Large gray arrows: non-radiative transition between states. Small gray arrows: vibrational relaxation to lowest vibrational state ($\sim 1 \times 10^{-11}$ s).

The transitions between electronic states of fluorescent molecules are illustrated in Jablonski diagrams. Here, the energetic states of molecules are represented as potential wells, which are populated by distinct vibrational states and electronic transitions between energetic states are represented as vertical lines. An example for a simplified Jablonski diagram showing transitions between S_0, S_1 and T_1 and the time scales associated with individual transitions is shown in Fig. 2.1. The Quantum yield (QY) of a fluorescent substance in a given environment is a measure for the probability of excited state relaxation via photon emission. It is defined as the ratio of absorbed and emitted photons:

$$QY = \frac{\# emitted\, photons}{\# absorbed\, photons} \tag{2.1}$$

As a consequence, a high QY is reached, if the rate of radiative relaxation from S_1 to S_0 via emission of a photon is favored over alternative, non-radiative transitions. In addition to the quantum yield, the molar extinction coefficient, η, which measures the probability of a molecule to absorb photons of a given wavelength, is typically used to characterize fluorescent labels. The molecular brightness, $QY\eta$ is helpful to compare fluorophores with respect to their expected signal strength in fluorescence-based assays. It is important to note that the properties of fluorescent molecules can strongly depend on the specific environment in which the molecules reside.

Due to its high stability compared to the S_1, the triplet state T_1 often serves as the starting point for reactions with other molecules. Since these reactions can have a large impact on the fluorescence signal observed from a fluorescent molecule, some of the relevant reactions and how they can be controlled will be discussed in the following.

It has been known for decades that the presence of molecular oxygen during excitation of fluorescent molecules negatively influences their stability and leads to irreversible bleaching[24]. This observed macroscopic behavior can be explained by the high reactivity of electronically excited triplet states of fluorophores with molecular oxygen ($^3O_2$). Reaction between these two reaction partners can lead to formation of highly reactive singlet oxygen ($^1O_2$) and superoxide anions, which in turn was shown to react with and ultimately destroy fluorescent molecules[25]. One simple solution to this problem is removal of molecular oxygen from the environment of fluorophores during experiments[26]. This can be efficiently achieved by addition of an enzymatic system that consumes oxygen in a catalytic reaction[27,28]. While removal of oxygen generally leads to reduced photobleaching, it was also observed that the lack of molecular oxygen as reaction partner to facilitate the transition from the triplet to the ground state increases the probability of fluorophores residing in dark states[29]. Alternative triplet state quenchers (TSQs) such as β-mercaptoethanol, Trolox or ascorbic acid can be used to promote the triplet state - ground state transition, but tend to exhibit varying efficiencies in dependence of the fluorescent molecule used[30,31].

A more consistent performance was observed using a combination of reducing and oxidizing agents as radical scavengers (ROXS)[32]. Supplying both types of reaction partners (reducing and oxidizing) is important as fluorophores may exhibit either reductive or oxidative reactive behavior once in their triplet state. Those reactions then lead to the formation of radicals and ultimately the destruction of the fluorophore (i.e. bleaching). Presence of the complementary reducing or oxidizing agent can prevent this reactive chain by rapidly reacting with the formed radical to return the fluorophore to its ground state S_0 (illustrated in Fig. 2.2). If used in combination with enzymatic oxygen removal the use of a Reducing and oxidizing system (ROXS) consisting of ascorbic acid as reducing and methyl viologen as reducing agent led to increased photostability, while at the same time effectively depopulating the triplet state.

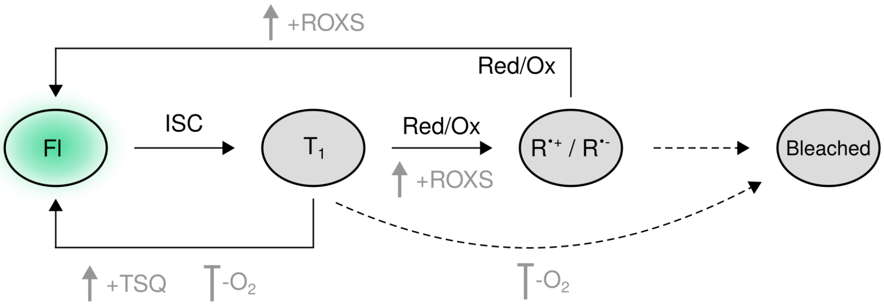

Figure 2.2 Influence of photostabilizing agents on fluorophores. Schematic overview of fluorophore state transitions between fluorescent (Fl), triplet (T_1), Radical cation/anion ($R^{\bullet}+/R^{\bullet}-$) and bleached state. State transitions indicated by labeled arrows. Red/Ox: Reduction or oxidion reaction. Intersystem crossing (ISC): intersystem crossing. Influence of oxygen removal (-O_2), addition of ROXS reagents (+ROXS) and addition of triplet state quenchers (+TSQ) indicated by symbols. Dashed arrows: Irreversible bleaching via diverse reactions.

Optical microscopy has been used to study biology ever since the first developments of single-lens and compound microscopes to observe microscopic specimen otherwise invisible to the human eye[33]. Developments in optical microscopy have led to advances across different biological fields and light microscopy continues to play an important role in biomedical research until today[34]. Modern optical systems offer minimal aberration and high light collection efficiencies. The combination of those with fluorescent labels and spectral detection gave rise to fluorescence microscopy. In contrast to a conventional light microscope, a fluorescence microscope makes use of the properties of fluorescence introduced above to selectively report on the presence of fluorescent molecules. This is achieved by restricting the spectrum of detected light to a narrow window matching the emission spectrum of the fluorescent marker, while at the same time preventing excitation light from reaching the detector. As a consequence, the sensitivity with which a fluorescence microscope is able to detect fluorescent molecules within a sample directly depends on its ability to collect and detect emitted light. In the past years, objective lenses with high light collection efficiency and detectors with quantum efficiencies >90 % have enabled fluorescence microscopists to approach the ultimate limit of sensitivity by detection of individual photons emitted from fluorescent molecules[35].

Despite these capabilities, fluorescence microscopy is still limited in its ability to resolve objects due to the diffraction of light. According to Abbe[36], even for a perfect imaging system free of aberrations, a fundamental resolution limit depending on the wavelength of light, λ and the numerical aperture (NA) of the objective applies:

$$d = \frac{\lambda}{2NA} \tag{2.2}$$

This means that objects located closer than the distance d along the lateral axis will not be distinguishable in images acquired with the optical system. For fluorescence microscopy using visible light for excitation and high numerical aperture objectives, the Abbe limit is in the range of 200-300 in the lateral dimension. For objects smaller than this diffraction limit, the signal of an object in the image plane of a microscope will appear as a diffraction-limited pattern, called the Point-spread function (PSF)[37]. Since the size and shape of the PSF is dictated by the properties of the optical system, it is possible to model the distribution of signal across a PSF to infer the position of the light-emitting object with precision beyond the resolution limit[38]. Today, this approach is being routinely used in single-molecule fluorescence microscopy, where the signal emitted by individual fluorophores is modeled with a two-dimensional Gaussian function to determine its position with best possible accuracy[39]. Additional information such as the shape, the width and the amplitude of the modeled Gaussian function can be used to assess additional properties of individual fluorophores[40].

Based on this simple idea, two classes of techniques with major impact on biological research have been developed. In a scenario, where fluorescent emitters are present at a density low enough such that PSFs of individual emitters do not overlap, the position of emitters can be tracked over time by repeatedly localizing them in a time series of individual images[41]. The second approach known as Single-molecule localization microscopy (SMLM) assumes a situation where the density of fluorophores is too high to identify individual PSFs. Here, the majority of fluorophores is first forced into a non-fluorescent state and subsequently only a small subset of fluorophores is transferred into a fluorescent state. The position of active emitters can then be determined before they are irreversibly photobleached (see

Fig. 2.2 above). By repeating this process many times, the exact position of fluorophores in the sample is accumulated and can be used to render an image with improved resolution. Depending on the process used to transfer fluorophores from their non-fluorescent state these techniques are known as photoactivation localization microscopy (PALM)**Betzig:Hess:2006** or stochastic optical reconstruction microscopy (STORM)[42]. For both of these techniques the localization precision, that is the accuracy with which the position of a fluorophore can be determined, is directly connected to the amount of photons collected from individual emitters in an image and therefore depends on the molecular properties of the fluorophores themselves[43].

2.2 Fluorescent labels

Fluorescent labeling is a prerequisite for any type of fluorescence microscopy. An extensive toolbox of fluorescent labels has been developed in the past and is available to meet the demands posed by different experimental systems, targets and fluorescence microscopy techniques[18,44,45]. In the following, I will briefly introduce different types of fluorescent labels with a focus on molecule-specific labeling of POI. I will then introduce self-labeling protein tags, which were used as the main labeling approach throughout this study in more detail. Finally, I will explain how labeling approaches can be calibrated to determine the number of labels or fluorophores attached to their target.

In its simplest form, fluorescent labels can be considered as a bifunctional system, where one part of the label mediates binding to the target and the other part of the molecule carries one or more fluorophores. While the binder part of fluorescent labels is critical to achieve specific and efficient labeling of the target, an ideal binder should be as small as possible to minimize potential impacts of labeling onto the target, while at the same time maintaining high specificity and affinity towards the target. The fluorescent part of the label serves to enable efficient detection of labels via emission of a fluorescent signal. In general a high molecular brightness (see above), photostability and low tendency to undergo dark-state transitions are considered favorable properties. Beyond the specific demands on the binder and fluorophore part of a label, general properties such as the overall size of labels, their compatibility with biological samples and their versatility with respect to possible targets are important criteria when choosing a label. It is important to note that depending on the imaging technique, these general demands on fluorescent labels may be extended towards specific properties. Fluorescent markers for labeling POIs can be broadly divided into two classes based on the means by which they are introduced into the sample. Affinity labels, which do not require modification of the sample prior to labeling and genetically-encoded labels, which have to be introduced into the biological system by expression of a transgene.

Affinity labels include small-molecule reagents with high specificity towards individual target, such as filamentous actin[46] or microtubules[47]. While these labels have a small size, can be labeled with a defined number of small organic fluorophores and often bind their target with high affinity, they do not offer a generalizable approach where arbitrary POI can be targeted.

Immunoreagents are a second group of popular affinity labels used extensively in biomedical research[48,49]. Same as small-molecule labels, immunoreagents require the development of a specific reagent (e.g. an antibody) for each desired target, typically by immunization of animals and harvesting of generated antibodies from blood samples. The most popular immunoreagent currently in use are IgG-type antibodies (Fig. 2.3). With a size of approximately 160 kDa, they are relatively large and typically require fixation and permeabilization to allow for access to intracellular targets. Antibodies are typically labeled with amine or thiol-reactive small organic fluorophores[50]. Since antibodies typically exhibit multiple reactive sites for attachment of such fluorophores, a distribution of fluorophore numbers attached to individual antibodies is often observed[51]. The observed variability in fluorophore numbers attached to target molecules via antibodies is further amplified, if polyclonal mixtures of different antibodies against a target are used. The same holds true if the target is labeled via indirect immunolabeling, where the primary antibody binding the target is not directly labeled, but a mix of fluorescently-labeled polyclonal secondary antibodies directed against the primary antibody is used. To reduce the size and the variability in fluorophore attachment, antibody fragments, which are obtained by protease digestion of antibodies can be used[52]. As it was shown that conjugation of fluorophores can influence the affinity of antibodies towards their targets[53], the functionality and affinity of antibodies should always be confirmed after conjugation. In summary, their size, the potential for steric hindrance upon binding of multiple antibodies to target molecules in close proximity and, limited accessibility towards target molecules in dense structures complicate the use of antibodies for quantitative imaging.

To reduce the variability often encountered with IgG antibodies, single-domain antibodies with a size of 15 kDa, often called nanobodies, derived from the variable region of Camelidae immunoglobulins can be used[54]. After establishing a new nanobody, they can be expressed in prokaryotic cells and labeled by site-specific modification, which largely reduces the variability in fluorophore number observed for IgG-type antibodies[55]. However, the same requirement of immunization of animals for raising novel nanobodies persists[56].

A third group of affinity labels consists of synthetic labels consisting of RNA (aptamers)[57] or peptides (affimers)[58] which can be generated by means of in vitro selection[59,60]. In addition, aptamers and affimers typically are small and, due to their simpler molecular structure enable stoichiometric labeling with a defined number of fluorophores[61,62].

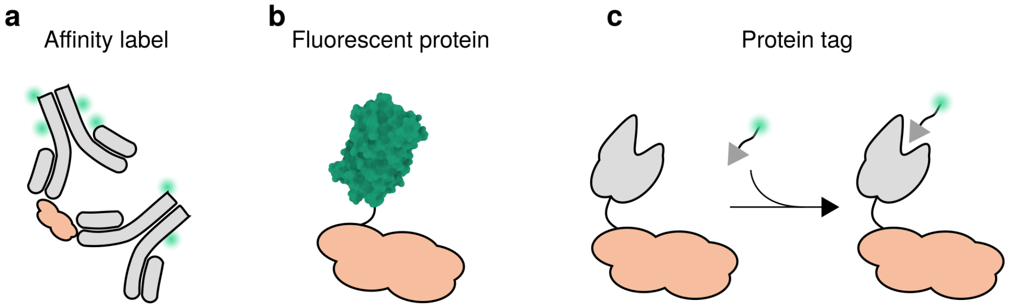

Figure 2.3 Overview of fluorescent labeling schemes. Schematic overview of different approaches to label target proteins (orange) with labels (gray) to attach fluorophores (green) to target proteins. Schemes not to scale. **a,** Labeling with fluorophore-conjugated polyclonal primary antibodies. **b,** Labeling by genetic fusion with a fluorescent protein. **c,** Two-step labeling via genetic fusion of protein tag and subsequent binding of fluorescent ligand to tag.

Genetically-encoded labels constitute the second class of widely-used labels for fluorescence microscopy. In contrast to affinity labels, genetically-encoded labels require introduction of a transgene into the host cells prior to imaging. While different simple and efficient methods for transient expression of transgenes in eukaryotic cells are available, the development of genome-editing technologies such as CRISPR/Cas9 now enables targeted introduction of genetically-encoded labels without perturbation of endogenous gene expression[63].

Fluorescent proteins are arguably the most prominent genetically encoded label and a large selection of fluorescent proteins with different properties is available[45,64]. In addition to fluorescent proteins, alternative genetically-encoded labels for fluorescence microscopy have been developed[65]. In contrast to fluorescent proteins, many of these protein tags consist of two complementary components. One component is genetically-fused to the POI and typically not fluorescent itself. The second component is either expressed as separate protein or introduced during an additional labeling reaction (Fig. 2.3c). This two-component scheme has the advantage that protein-specific labeling with small organic fluorophores can be achieved in a live-cell context. At the same time, two-component protein labels can easily be attached to different POI and often allow for labeling with different fluorophores due to their modular structure[66].

Early examples of two-component tags include FlAsH tag[67] and eDHFR[68]. FlAsH tag consists of a 12 amino acid peptide containing a tetracysteine motif and a bis-arsenical fluorescein probe which binds to the tetracysteine motif. While labeling with FlAsH tag is limited to a specific fluorescent ligand, eDHFR is able to bind to different trimethoprim (TMP) derivates as ligands and exhibits fast ligand binding kinetics, but does bind its ligands in a covalent reaction. An modified version of eDHFR capable of covalent ligand binding named TMP-Tag was reported[69].

A relatively recent class of labels, are self-labeling proteins, which combine a set of favorable properties for advanced fluorescence microscopy approaches. Self-labeling proteins covalently bind their substrates through a single active site in their catalytic center, they exhibit fast kinetics and were shown to be compatible with a wide range of ligands[70]. The first reported self-labeling protein tag, SNAP-tag, is based on the human DNA repair enzyme O^6-alkylguanin-DNA alkyltransferase (hAGT, Fig. 2.4a) and binds O6-Benzylguanine (BG) or benzylchloropyrimidne (CP) derivatives as ligands (Fig. 2.4b). Covalent binding occurs via the cysteine residue Cys145 in the catalytic center of the protein under release of guanine (Fig. 2.4c). The first version of SNAP-tag, SNAP26m contained 10 mutations compared to hAGT and was shown to be able to bind Background (BG)-ligands after expression in living eukaryotic cells[71]. SNAP_f-tag was developed as an improved version of SNAP-tag by introducing another 19 mutations[72] and was shown to exhibit >10-fold increased labeling kinetics**Sun:CorreaJr:2011**. In addition, CLIP-tag was developed as a closely-related self-labeling protein tag binding to O2-benzylcytosine derivatives as ligand[73].

Based on the prokaryotic dehalogenase DhaA isolated from *Rhodococcus sp.*, HaloTag (Fig. 2.4d) was developed as an orthogonal self-labeling protein tag to SNAP_f-tag and CLIP-tag[74]. HaloTag reacts with choloralkane ligands[75] to which a wide variety of functional groups can be conjugated (Fig. 2.4e). HaloTag binds to its ligands via formation of an ester bond between aspartate asp106 and the hydrocarbon chain of the ligand followed by release of a chloride ion (Fig. 2.4f). A recent comprehensive study compared the labeling kinetics for SNAP_f-tag and HaloTag and reported apparent rate constants for HaloTag labeling with the HaloTag ligand HTL-TMR of $>1 \times 10^7\ \text{M}^{-1}\,\text{s}^{-1}$, almost two order of magnitude higher than what coud be achieved with SNAP_f-tag using a comparable BG-TMR liand[70]. An improved variant of HaloTag, HaloTag7, was further optimized with respect to its stability and reactivity[76]. A second favorable property of HaloTag is that fluorescent ligands can interact tightly with the surface of HaloTag, which has enabled the development of fluorescent ligands, which exhibit fluorogenic properties upon binding to HaloTag through a shift in the equilibrium of a intramolecular spirocyclization reaction[76,77]. Recently, several mutants of HaloTag with additional improvements over HaloTag7 were reported[78,79].

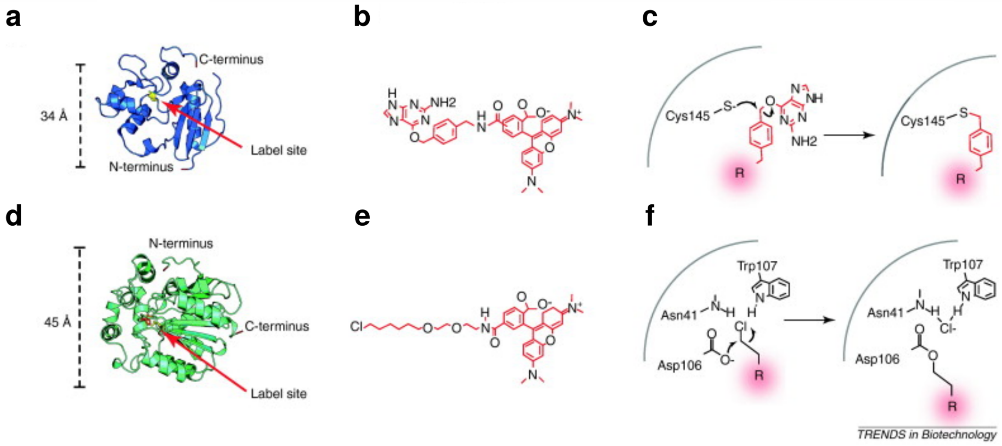

Figure 2.4 Protein structure, substrates and reaction mechanism for SNAP-tag and HaloTag. a-c Protein structure (a), structure of ligand BG-TMR (b) and reactive site (c) for SNAP-tag. **d-f** Protein structure (d), structure of ligand HaloTag ligand (HTL)-TMR (e) and reactive site (f) for HaloTag. Figure reproduced under license from[80].

In addition to the protein tags introduced above, a series of recently developed peptide tags represent a promising alternative for quantitative fluorescence microscopy approaches. Peptide tags feature a small size (12-15 amino acids) of the peptide that has to be genetically fused to the POI and high-affinity binders based, for example, on nanobodies are available[81,82]. Finally, Unnatural amino acids can be considered the smallest possible protein label based on genetic fusion[83]. In combination with bioorthogonal chemistry for attachment of fluorescent labels, UAAs represent a promising technique for quantitative approaches since they enable labeling of POIs with minimal perturbation, while at the same time enabling the attachment of small organic fluorophores[84]. It was recently reported that by using an inverse electron-demand Diels Alder reaction for conjugation of tetrazines to dienophiles, probes with fluorogenic properties and high reaction kinetics were obtained[85–87].

In addition to the general properties of fluorescent labels discussed above, the suitability of fluorescent labeling approaches for quantitative studies of protein numbers depends on three additional criteria. First, the target binding efficiency of a given labeling approach directly influences our ability to draw conclusions about the number and distribution of target molecules based on the presence of fluorescent labels. Second, nonspecific binding of the fluorescent label to samples will limit our ability to infer target molecule numbers from the number of fluorescent labels in a sample. Third, labels for quantitative fluorescence microscopy should exhibit a defined and narrow distribution of fluorophores per label. For an ideal label, the target binding efficiency should thus be as high as possible, while the tendency for non-specific attachment should be as low as possible. For many counting approaches, stoichiometric labeling, i.e. a 1:1 relation between label and fluorophore is desirable. To determine the quality of a given labeling approach, it is therefore necessary to characterize all three of these properties in dedicated assays.

The Degree of labeling (DOL), refers to the number of fluorophores attached to a label on average an dis typically measured in vitro by absorption spectroscopy using published information about fluorophore spectra and brightness[88]. However, reports showed that fluorophore conjugation to antibodies resulted in changes of fluorophore brightness, which can lead to biased DOL values if not corrected for[53]. In addition, since the DOL only reports on the ration of fluorophores and labels across an ensemble of label molecules, heterogeneity in the distribution of fluorophores remains undetected, which can lead to bias in subsequent quantitative measurements. This an be addressed by measuring the label number distribution instead of the average DOL using an approach with single-molecule sensitivity[51].

The labeling efficiencies refers to the fraction of POI caries a fluorescent marker. The labeling efficiency can be determined with different methods ranging from semi-quantitative in vitro approaches based on immunoblotting[71] to more quantitative approaches based on mass-spectrometry[89]. In situ assays based on fluorescence microscopy are advantageous as they report the labeling efficiency for conditions that are comparable to the ones used in subsequent quantitative fluorescence microscopy measurements[90,91]. This allows to determine the effective labeling efficiency, i.e. the fraction of labels that can be successfully detected with a given experimental approach[92]. For SNAP_f-tag, labeling efficiencies ranging from 30 % up to >90 % were reported for different ligands and using different methods[93]. For HaloTag, generally slightly lower labeling efficiencies around 30-40 % were reported[92,94].

For evaluating non-specific binding of fluorescent labels, it is especially important to ensure that the conditions used for testing the labeling specificity match the conditions that will later be used for measurements targeting POI. As part of the evaluation of a fluorescent labeling approach, non-specific label deposition should always be checked for to allow correct interpretation of fluorescence microscopy data. This is typically achieved by generating a negative control sample which lacks specific bindings sites and which is then treated and imaged in the same way as the sample[95].

2.3 Emitter Counting Fluorescence Microscopy

Light microscopy is limited in resolution due to the diffraction of light. Different methods have been introduced to improve the spatial resolution of fluorescence microscopy beyond the diffraction limit**Sauer:Heilemann:2017**, **Sahl:Jakobs:2017**, [43]. Still, the resolution of these methods is not yet sufficient to enable unambiguous identification and localization of all individual fluorophores in a given sample. Therefore, counting of fluorescently labeled molecules by direct enumeration of fluorophore signals in images is not (yet) possible. Instead, it is therefore necessary to make use of additional information contained in the fluorescence signal to determine the number of fluorophores that contributes to the measured signal.

Different methods to convert fluorescence signals into fluorophore numbers have been developed[96,97]. The first class of techniques is based on fluorescence fluctuation approaches and provides estimates of the concentration of fluorescently labeled POI by measuring intensity fluctuations caused by molecules diffusing within the sample[98–100]. Although these methods can be used to determine the stoichiometry of protein complexes, they are limited in their ability to address stationary complexes and in their ability to provide information about the distribution of stoichiometries for heterogeneous complexes. For this reason, they will not be discussed further in the context of this study. Comprehensive reviews covering this class of techniques are available[101–103].

A second class of methods provides information about the absolute abundance of labeled molecules in defined structures, which can be identified and localized based on their signals observable in images[19]. These methods often are able to report measured fluorophore numbers for individual structures and thereby give access to the distribution of fluorophore numbers across a population of individual structures. I will refer to this set of methods as emitter counting microscopy techniques in the course of this study.

In the following, I will provide an overview of the different concepts for emitter counting microscopy, provide selected application examples and discuss the advantages and disadvantages of individual concepts.

One straightforward way to obtain fluorophore number estimates, is to compare the fluorescence intensity recorded from a structure to the intensity emitted from a single fluorophore or a structure labeled with a known number of fluorophoresFig. 2.5. This emitter counting approach, also known as Intensity-based counting (IBC), has been applied as early as 1986[104] and poses little demands on the type of fluorescence microscope data is acquired on or on fluorescent markers used for labeling. IBC can either be used to assign fluorophore numbers to individual clusters[105] or to assess the stoichiometry of an ensemble of molecules[106] and different implementations of IBC have been demonstrated[107,108]

The accessible counting range of IBC for assigning fluorophore numbers to individual signals is limited by the increasing uncertainty of fluorophore number estimates and was shown to not exceed 4[105]. Stoichiometry estimates for ensembles of molecules are not affected by this limitation and higher number of fluorophore number estimates were reported following the ensemble approach to IBC. For in vitro samples, where the density of individual complexes can be adjusted and complexes can be immobilized directly on the surface of coverslips, IBC was found to have a high precision. Experimental validations of IBC with DNA origami demonstrated an accessible counting range of up to 36 counted fluorophores per cluster[109]. Another example where IBC was applied for characterization of protein complexes in cellular samples was recently reported by Finan et al.Finan:Heilemann:2015. Here, protein complexes with up to 32 protein copies were characterized by comparison of intensity distributions acquired on many individual complexes with the average intensity of defined protein oligomers with known stoichiometry.

The main limitation of IBC results from the underlying assumption that the unitary brightness for the calibration probe (individual fluorophores or complexes with known stoichiometry) has to be identical to the unitary brightness of fluorophores in the target complex. In practice, this assumption is often violated due to variable excitation intensities across the field of view, e.g. due to uneven illumination or due to Total internal reflection (TIRF) illumination. Also, fluorophores themselves can exhibit different behavior depending on their local environment. For example, it was reported that small organic fluorophores exhibit changes in photophysical properties after conjugation to proteins[110] and upon changes in pH[111]. In addition, the need for a calibration sample poses high demands on minimizing sample-to-sample variability as minor changes in sample preparation or imaging conditions can lead to inaccurate emitter number estimates upon comparing sample and reference intensities.

Irreversible photobleaching of individual fluorophores was initially used to demonstrate the successful detection of individual fluorophores by fluorescence microscopy[112]. The same principle was later applied to determine the number of fluorophores contained in a

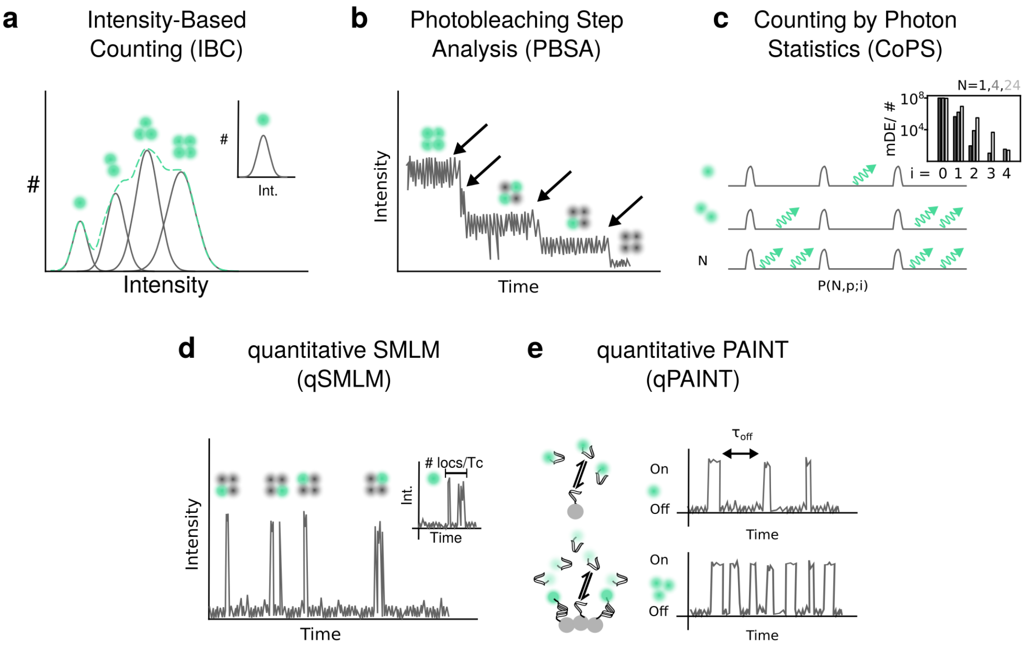

Figure 2.5 Overview of emitter counting microscopy techniques. a, Intensity-based counting by comparison with calibration standard (inset). **b,** Emitter counting by identification of individual photobleaching steps (arrows) in intensity traces (gray). **c,** Principle of antibunching utilized in Counting by photon statistics (CoPS). Inset: Multi-detection event distribution for different emitter numbers (1,4,24). **d,** Quantitative single-molecule localization analysis. Inset: Calibration standard to correct for average number of localizations per label. **e,** Quantitative Points accumulation for imaging in nanoscale topography (PAINT). Binding site numbers are measured from repeated binding of imager strands (left) and dark time distribution without imager strand bound to target complex (right). Figure modified after[19].

complex[113]. For this, a sample is exposed to continuous, high-intensity illumination and the intensity from a Region of interest (ROI) is recorded over time. Plotting the intensity values of those ROIs over time generates intensity traces, which are used for the subsequent analysis. Step-wise losses in fluorescence intensity caused by irreversible photobleaching of individual fluorophores are then identified in the intensity trace and counted to determine the total number of fluorophores present at the start of the measurement (Fig. 2.5b). Unlike IBC, Photobleaching step analysis (PBSA) was shown to achieve a higher accessible counting range, but ultimately is also limited by noise in cases where many fluorophores contribute to an fluorescence intensity trace. In addition, the likelihood for missed steps due to multiple photobleaching events occurring within a single time point increases exponentially with increasing fluorophore numbers. Such a behavior can be minimized by adjusting imaging rates and exposure times. Similar to IBC, PBSA is compatible with a wide range of microscopy techniques for data acquisition and was shown to be compatible with different fluorescent markers[114,115]. However, fluorophores, which exhibit transient dark-state transitions will complicate trace interpretation and ideally would be avoided when choosing a fluorescent marker for PBSA. Since PBSA relies solely on the identification of step-like intensity changes in individual traces, it is able to provide fluorophore number estimates for individual complexes and is calibration-free.

In its initial form, PBSA was non-automated and relied on manual classification of photobleaching traces to assign perceived numbers of step-wise intensity changes[113,116]. Such approaches were successfully used for biological targets with stoichiometries of <10 per complex[117]. A first improvement of PBSA approaches was achieved by filtering of intensity traces to remove noise and to facilitate manual interpretation[118]. Here, for example Chung-Kennedy filtering was used to preserve step-like features while reducing overall noise[114].

Progress towards automated analysis of photobleaching traces was made by developing approaches based on on hidden Markov modeling[119] or change-point analysis[120] to assess the oligomeric state of G protein-coupled receptors in vitro with counting of up to six subunits[121]. This approach was validated using Enhanced green fluorescent protein (EGFP)-multimers and was found to be highly dependent on the achieved Signal-to-noise ratio (SNR) during image acquisition.

A more generalized approach to PBSA addressed the need for more sophisticated extraction of intensity traces from cellular samples, where high densities of target complexes and cellular background otherwise complicate trace interpretation. This was combined with Chung-Kennedy filtering for noise reduction and a custom approach for identification of photobleaching steps in filtered traces. The established pipeline was validated on the pentameric transmembrane receptors GluK2 and GlyRα1 labeled with EGFP[122].

Recently, a set of fully automated step detection algorithms based on Bayesian approaches was established and reported to achieve extended accessible counting ranges of more than 30 fluorophores in individual traces[123,124]. While such appraoches clearly represent a significant step in the ability of PBSA-based emitter counting microscopy to target large intracellular complexes, experimental validation on complexes with stoichiometries >30 has not been reported to date. In addition, due to the high complexity of these algorithms computational cost becomes an relevant factor with run times of 17 h per trace having been reported for analysis of a trace containing 50 photobleaching steps[123].

A third class of emitter counting approaches is based on the principle of photon antibunching, which dictates that the probability for detecting more than one photon from a fluorophore approaches zero for a time window approaching a width of zero[125]. In contrast to IBC and PBSA, emitter counting based on photon antibunching requires specialized microscopy setups equipped with single-photon sensitive detectors which are typically used in conjunction with confocalized detection. Photon antibunching can be used to infer the number of fluorophores by recording coincident photon detection events from simultaneously excited fluorophores in a diffraction-limited structure[126] (Fig. 2.5c). By employing a detector geometry where the collected fluorescence signal from a sample is equally distributed onto four single-photon avalanche diodes, multiple photon detection events (mDE) with up to four coincident photons can be recorded (Fig. 2.5c, inset). The distribution of mDEs detected after excitation with pulsed lasers with pulse widths shorter than the fluorescence life time of the fluorophores used (typical pulse widths: $\sim$100 fs), then allows to infer the number of independent fluorophores present in the excited volume[127]. However, in contrast to PBSA and all other emitter counting microscopy approaches except for IBC, CoPS is able to provide time-resolved fluorophore number estimates. Here, the achievable time resolution and counting precision depends on the molecular brightness of the employed fluorophores and sub-second time-resolution was previously reported[128,129].

CoPS was shown to achieve an experimentally validated counting range of at least 30 fluorophores[130], while simulations showed that accessible counting ranges of up to 50 fluorophores could be possible[131]. By combining CoPS with the super-resolution technique Stimulated emission depletion (STED), quantitative information on emitter numbers across a sample were obtained, while providing spatial information with improved resolution[132].

Following the same concept of combining improved spatial resolution with quantitative information about fluorophore numbers emitter counting microscopy methods based on SMLM have been described. quantitative Single-molecule localization microscopy (qSMLM) is based on the concept that individual localizations used to reconstruct super-resolved images in SMLM contain information about the underlying number of blinking fluorophores[133]. By correcting the number of localizations recorded from a given structure to the number of localizations recorded on a reference standard where the number of fluorophores per target is known, the underlying target copy number in the structure can be determined (Fig. 2.5d). This concept has been extensively used in combination with photoactivatable fluorophores[134–137] and small organic fluorophores which can be shelved in transient dark stats[138,139].

While the basic idea behind qSMLM is intriguingly simple, the implementation of qSMLM for calibration-free counting can be challenging and is currently limited to accessible counting range $\ll$10[139,140]. In contrast, counting of target molecules by calibration with reference standards was applied to structures containing $\gg$100 fluorophores[92,138]. In this context, it is important to note that qSMLM can be readily applied to structures with arbitrary shapes.

An interesting extension of qSMLM is a recently developed technique named quantitative Points accumulation for imaging in nanoscale topography (qPAINT). qPAINT relies on the concept of transient binding and unbinding of fluorescently labeled DNA oligonucleotides to complementary oligonucleotides. Here, localizations of fluorescent signals are generated from signals emitted by transiently bound oligonucleotides[141]. This effectively uncouples the generation of blinking signals required for SMLM from the properties of individual fluorophores and thereby enables SMLM with high localization precision and label densities. The fact that the binding and unbinding of oligonucleotides is a well-defined process with simple kinetics, the frequency of binding events contains information about the number of binding sites present in a structure. This relation is used by qPAINT to deduce the number of binding sites from experimentally measured distributions of dark times and by comparison with a reference standard with a known number of binding sites[142]. qPAINT was shown to exhibit a large accessible counting range and excellent precision for in vitro measurements with DNA origami. Recently experimental factors influencing the precision of qPAINT measurements have been described[143]. qPAINT has been applied to measure protein copy numbers in different biological systems[144–146] and a calibration-free approach for qPAINT on low copy number targets was demonstrated recently[147]

2.4 Standards for Quantitative Fluorescence Microscopy

The need for validation and standardization in quantitative fluorescence microscopy has given rise to a number of approaches which aim at improving the reproducibility of microscopy techniques[148,149]. Alongside efforts to standardize data formats[150,151], reporting of experimental parameters[152] and techniques for characterization of optical systems[153,154], reference standards have emerged as important tool for benchmarking fluorescence microscopy techniques[155,156]. In the following, I will introduce different reference structures with dimensions below the diffraction limit and which are therefore of particular use for emitter counting microscopy.

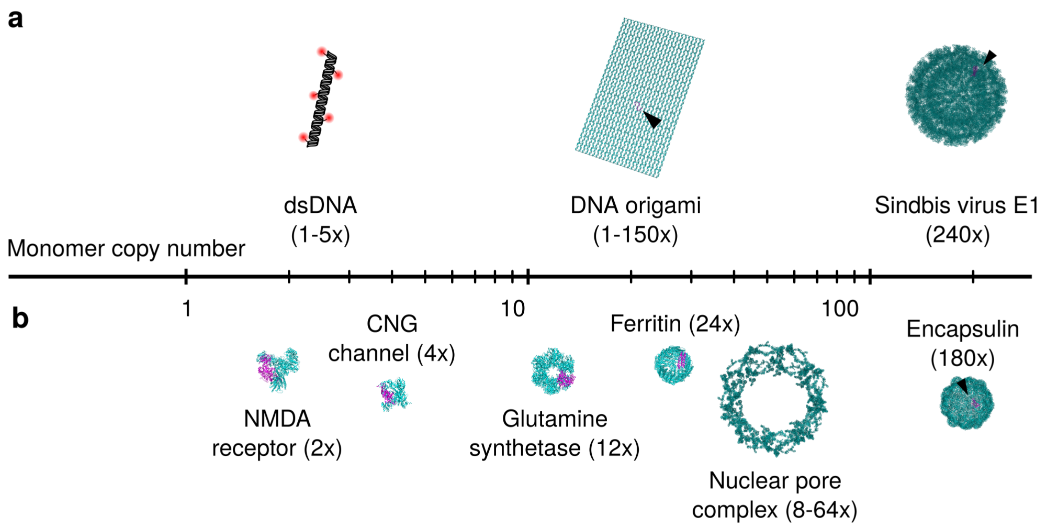

Figure 2.6 Selected reference standards for protein copy number measurements. Standards for emitter number measurements. Protein oligomers and DNA-based probes allow to attach a defined number of fluorophores to complexes with variable size. **a,** Standards for in vitro studies. DNA oligonucleotides, DNA origami and viruses can be used as reference standards with up to 240 fluorophores attached to a individual structures <100 nm. **b,** In-cell reference standards. Different trans-membrane and soluble protein complexes with monomer copy numbers ranging from 2–180 have been described in the literature. Protein structures were rendered using Chimera[157] based on publicly available data deposited in PDBe. Accession numbers: 5iou (NMDAR), 5h3o (CNG), 1fpy (Glutamine synthetase), 1eum (Ferritin), 4pt2 (Encapsulin), 3j0f (Sindbis virus), 5a9q (Nuclear pore complex). The structure of a rectangular DNA origami was predicted using CanDo[158] and rendered using Chimera.

Reference standards for emitter counting microscopy can be distinguished into two classes based on their area of applicability. The first class of reference standards was developed for use as isolated particles and their application is often restricted to in vitro studies. Examples for this class of standards include DNA oligomers[159], DNA nanostructures[109], viral particles[160] and hybrid DNA-protein nanostructures[161] (Fig. 2.6a). Among these examples, DNA origami have emerged as widely used reference samples due to their flexible design, ease of use and high stability under various conditions[162]. DNA-based nanostructures are typically covalently labeled with small molecule fluorophores and require purification after labeling[163]. Approaches to attach fluorescent proteins or other protein-based fluorescent labels to DNA nanostructures have been reported[161]. Although DNA origami are highly popular as reference standards, limited labeling efficiencies[164] and an influence of attached fluorophores on the origami itself were reported[165].

A second class of reference standards was established to meet the need for standards that can be expressed inside eukaryotic cells. This class of standards consists of different protein types, which were found to assemble into stable macromolecular complexes. While some of these protein-based reference standards occur naturally in eukaryotic cells and exert essential cellular functions, other protein complexes are artificially introduced into eukaryotic cells. Examples for this class of reference standards include different types of transmembrane proteins[113,166,167], oligomeric enzyme complexes[168,169], the nuclear pore complex[92] and virus-like particles[169,170] (Fig. 2.6b). Labeling of in-cell reference standards is typically achieved with the same labels as for labeling of the POIs (see section 2.2). In contrast to reference standards for in vitro studies, the properties of in-cell standards can be influenced by regulatory processes that affect the complex stoichiometry within the host cell. Therefore, additional criteria beyond the stoichiometry of a complex such as the efficiency of complex formation in cells, density and subcellular localization are important criteria when choosing an in-cell reference standard for a given application.

2.5 Scope of this study

The aim of this study was to establish an experimental framework for counting protein copies in macromolecular complexes by quantitative fluorescence microscopy. To relate the observed fluorescence signal to protein copy numbers, two tasks have to be performed:

1. The fraction of POI successfully labeled with a fluorescent marker needs to be determined.

2. The number of fluorescent labels contributing to the fluorescent signal need to be inferred from the recorded fluorescence signal.

I addressed both of these challenges in the scope of two projects, which will be described in the following chapters of this study. In chapter 3, I will describe the development of a staining efficiency probe and a corresponding data analysis pipeline to determine the labeling efficiency achieved for protein labeling with self-labeling protein tags. This work was based on a previous proof-of-principle study performed in our lab, which established this staining efficiency probe as tool for identifying optimal labeling conditions for multi-color single-particle tracking and which provided first evidence that labeling efficiencies can be measured using the probe.

The first goal of my work was to identify and improve key steps in the data analysis pipeline to optimize the robustness of the pipeline. I then benchmarked the pipeline with simulations and experimental data to verify that unbiased labeling efficiencies could be measured across different experimental conditions. The second goal of this project was to translate the manual imaging assay into a semi-automated imaging workflow with increased throughput. Despite the widespread use of SNAP$_f$-tag and HaloTag, no systematic studies of labeling efficiencies that can be achieved with these tags have been reported to date. To fill this gap in our understanding of self-labeling protein tags, I used the optimized data acquisition and processing pipeline to investigate the performance of SNAP$_f$-tag and HaloTag under a wide range of different labeling conditions.

The second project, which will be presented in chapter 4, aimed at establishing an approach for in situ emitter counting microscopy without the need for specialized hardware to allow for easy implementation across laboratories. As described in section 2.3, photobleaching step analysis is a calibration-free technique that is able to provide information on the single-particle level and does not require specialized instrumentation beyond a widefield fluorescence microscope. It is therefore well suited for applications in cell biology and can be adapted to process raw data acquired on different types of fluorescence microscopes.

However, previously reported algorithms for photobleaching step analysis suffer from limitations due to a limited counting range, high computational demand required for trace interpretation or lack of validation with experimental data. The aim of this project was therefore to develop and validate a comprehensive framework for photobleaching step analysis with a special emphasis on computational performance to allow for studies measuring emitter numbers across many complexes per cell and many cells. In conjunction with developing a new algorithm for trace interpretation, a second goal was to identify experimental conditions suited for generating high-quality input data to maximize the accessible counting range of the developed approach. Finally, the developed algorithm was to be validated with simulated as well as experimental data to provide a robust estimate of the accessible counting range under relevant experimental conditions.

Part II

Results

Chapter 3

CALIBRATED LABELING WITH SELF-LABELING PROTEIN TAGS

To measure copy numbers for a POI with fluorescence microscopy, the fraction of POIs labeled with fluorescent markers needs to be determined. Self-labeling protein tags such as $\mathrm{SNAP_f}$-tag or HaloTag, offer the advantage that they are genetically fused to the POI and that they bind at most one fluorophore per tag (section 2.2). This comes at the cost that they require an additional labeling step where samples are incubated with fluorescent ligands. The possibility of non-specific deposition of fluorescent ligands in samples during labeling means that the labeling efficiency has to be determined while taking non-specific ligand deposition into account.

In the following chapter, I will first explain how labeling efficiencies can be measured on the single molecule level. I will then describe the development and validation of `DOLanalysis`, a pipeline for labeling efficiency measurements by single-molecule colocalization analysis. Finally, I will present experiments performed to assess the performance of $\mathrm{SNAP_f}$-tag and HaloTag under different experimental conditions with the goal of identifying optimal labeling conditions for robust protein counting by fluorescence microscopy.

Parts of the results described in this chapter were performed in collaboration with Felix Braun, who implemented `DOLanalysis` as a MATLAB framework for data analysis and helped with establishing simulations. Prior work was performed by Siegfried Haenselmann and where indicated, parts of the measurements were performed by Felix Braun, Siegfried Haenselmann and Wioleta Chmielewicz (all Herten Lab, Heidelberg University). A publication describing the results presented in this chapter is currently in preparation for submission. Felix Braun, Siegfried Haenselmann and me will be equal first authors of this publication.

3.1 Single-molecule labeling efficiency measurements

In section 2.2, I explained that fluorescent labeling approaches can be characterized with different experimental techniques, which provide limited amounts of information about the efficiency of a labeling approach on the ensemble level or about the relative efficiency of a labeling approach.

A direct way to measure the labeling efficiency and label number distribution on the single-molecule level is to assess the colocalization of the label with a second fluorescent marker, which serves as reference signal, that indicates the presence of a potential labeling site[91]. To implement a single molecule approach to measure labeling efficiency based on colocalization, five basic requirements have to be fulfilled:

1. The reference marker and the label under investigation need to be distinguishable based on their spectroscopic properties. This can, for example, be achieved by choosing fluorophores with different excitation or emission spectra.
2. Complexes between the POI, reference marker and the label being characterized have to be sufficiently stable to allow for detection of reference and marker signals.
3. The label being characterized and the reference marker have to be sufficiently bright for single-molecule detection.
4. The spacing between individual POIs within the sample has to be sufficiently large and their respective mobility sufficiently low to identify individual POI in fluorescence microscopy images and to enable assigning pairs between reference markers and labels being characterized.
5. Emitter detection for reference marker and label being characterized have to be specific, i.e. a majority of observed signals needs to originate from specifically introduced fluorophores.

By designing an artificial protein as Staining efficiency probe (SEP), we set out to meet the demands listed above. This SEP consists of a fluorescent protein, which serves as reference marker and two self-labeling protein tags, for which the labeling efficiency was to be calibrated (Fig. 3.1a). In the following, I will refer to signals detected from the fluorescent protein as reference marker or 'reference' and to signals emitted from ligands of self-labeling protein tags as target signals or 'target'. At low expression strength, signals from individual SEPs can be observed as diffraction-limited signals in a cell (Fig. 3.1b). The position of individual reference and target can be obtained by sub-pixel localization in separate spectral channels (Fig. 3.1c). This type of analysis also allows to distinguish between ligands specifically attached to a SEP and non-specific deposition of ligands in the sample (Fig. 3.1d).

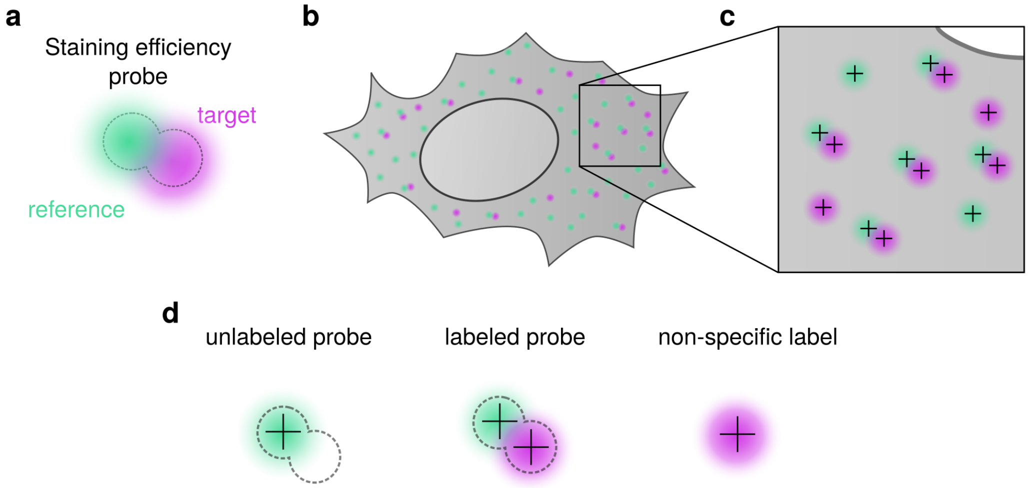

Figure 3.1 Labeling efficiency measurements by single-molecule colocalization analysis. a, SEP for single-molecule colocalization analysis between reference marker and target label. **b,** Probes are expressed at low density in cells **c,** Reference and target signals from individual probes can be detected and localized with high precision (crosses) in close spatial proximity. **d,** Signal from specific and non-specific labeling can be distinguished based on colocalization with a reference marker.

The labeling efficiency can then be obtained by computing the degree of colocalization between localized reference marker signals and localized signals from fluorescent ligands bound to the protein tag:

$$Degree\,of\,colocalization\,(DOC) = \frac{ref \cap target}{ref} \tag{3.1}$$

where $ref \cap target$ is the number of reference marker signals for which a signal from a protein tag ligand ($target$) was detected in close proximity and ref is the total number of reference marker signals. The DOL can be computed from the same data by determining the average number of signals from protein tag ligands (target) per reference marker position. In the case of labeling with protein tags, where tags are genetically fused to a target protein and each tag molecule can bind at most one fluorescent ligand molecule, the DOL is a direct measure of the labeling efficiency. In the following, the DOL will be reported as the fraction of labeled tags (ranging from 0 to 1), while the labeling efficiency will be reported as the percentage of labeled tags.

3.2 Prior developments

3.2.1 A staining efficiency probe for SNAP$_f$-tag and HaloTag

Based on the principles outline above, we created a SEP specifically designed to measure the labeling efficiencies for the self-labeling protein tags SNAP$_f$-tag and HaloTag with orange to red emitting fluorescent ligands. Detection of the reference signal in a separate spectral channel was enabled by using the green-emitting fluorescent protein EGFP as reference marker. Both tags were connected by an 25 amino acid α-helical linker with an estimated length of 57 Å[171] to reduce potential energy transfer via Foerster resonance energy transfer (FRET) between fluorophores conjugated to each tag. The SEP was targeted to the plasma membrane by addition of the N-terminal signal peptide of Lyn kinase, which is myristoylated by N-myristoyl transferase and consequently inserts into the inner leaflet of the plasma membrane[172]. At the C-terminus, a His$_6$-tag was added to allow for verification of full-length expression of the construct via immunoblotting. Due to the addition of EGFP as reference marker, we name this SEP EGFP-enhanced staining efficiency probe (gSEP). A schematic overview of all components of gSEP is shown in Fig. 3.2a. A second construct consisting of a EGFP targeted to the plasma membrane via N-terminal fusion to the Lyn signal peptide was generated to assess non-specific ligand deposition (Lyn-EGFP (LynG), Fig. 3.2b). Both constructs were stably expressed in mammalian cells by retroviral transduction with the Phoenix Ampho system[173].

TIRF microscopy of Huh7.5 cells stably transduced with gSEP or LynGs showed that the expression strength of both constructs was sufficiently low to identify individual diffraction-limited EGFP signals (Fig. 3.2c,d left) which were to be used as reference markers.

TIRF microscopy of Huh7.5 gSEP and LynG cells after labeling with the SNAP$_f$-tag ligand BG-SiR and the HaloTag ligand HTL-TMR further showed that both tags within gSEP could be labeled with their respective ligands in gSEP-expressing cells (Fig. 3.2c) while less signal was observed in the target (HTL-TMR and BG-SiR) channels in LynG-expressing cells (Fig. 3.2d, middle & right). A comparison with Huh7.5 gSEP cells not labeled with SNAP$_\text{f}$-tag or HaloTag ligands showed few autofluorescent signals indicating that imaging of tag ligands with high specificity was possible (Fig. 3.2e). In the following, I will refer to the EGFP channel as reference channel, while I will refer to the channels containing signals from protein tags as target channels.

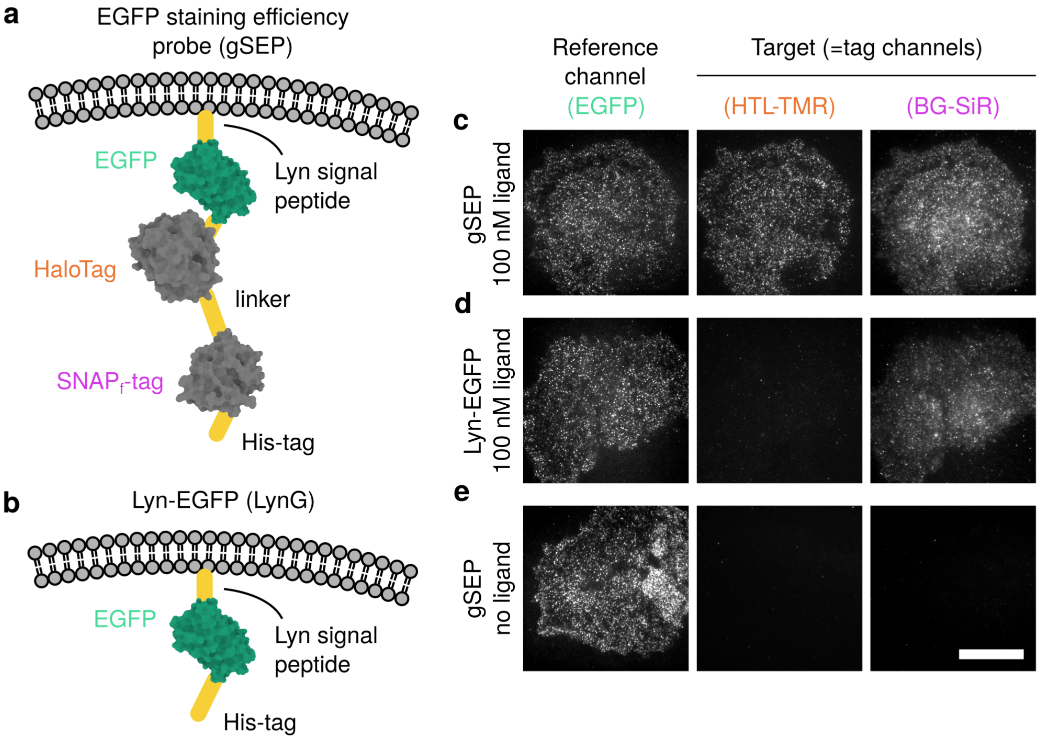

Figure 3.2 Principle and validation of gSEP activity. a,b Schematic structures of gSEP (a) and LynG (b). Molecular models of EGFP, SNAP$_\text{f}$-tag and HaloTag were obtained from PDB (accession numbers 1EMA, 6Y8P, 6Y7A). Scheme not to scale. **c-e,** Representative TIRF microscopy images of Huh7.5 gSEP or Huh7.5 LynG after labeling with HTL-SiR (500 nM for 30 min before fixation (c,d) or without labeling (e). Scale bar: 20 µm. Microscopy images in c-e recorded by Siegfried Haenselmann.

3.2.2 Labeling efficiency measurements with gSEP

From looking at images of cells expressing gSEP and labeled with fluorescent ligands, it is obvious that obtaining the degree of colocalization form raw images requires a computational pipeline for data analysis. For this purpose, Siegfried Haenselmann previously developed a pipeline to perform a single-molecule colocalization analysis which uses multi-channel images recorded from cells expressing gSEP or LynG as input and computes the degree of labeling based on emitters identified in individual image channels (**??**a). While the development of routines used in each step of the pipeline were described in detail before[174], I will summarize the underlying principles for each step below as they are essential to understand the subsequent improvements to this pipeline that I have developed as part of my thesis.

Input data: As described above, signals from reference and target fluorophores need to be distinguished to perform a single-molecule colocalization analysis. In our case, we used spectral information to distinguish between reference and target fluorophores and therefore obtained a multi-channel image set with 2-3 images (1 reference channel, 1-2 target channels) for each cell (Fig. 3.3b). Based on the reference channel image of each cell, a binary mask was computed as additional input for subsequent processing steps (Fig. 3.3c).

Emitter localization: Individual emitters in images have to be identified and localized with sub-pixel precision. Typically this is achieved by fitting a two dimensional Gaussian function to pre-selected sub-regions of images. In addition to the precise x,y coordinates of an emitter, this gives access to additional properties such as the signal strength (amplitude of Gaussian) and the signal width (sigma of Gaussian) for each emitter (Fig. 3.3d). Typical challenges during this step include preselection of sub-regions for fitting in noisy images and separation of partially overlapping emitters. Since Gaussian fitting for determining sub-pixel locations of emitters is routinely performed in SMLM and other single-molecule fluorescence microscopy approaches, many algorithms and software packages to address this task are available[175]. In the initial version of the gSEP data analysis approach, emitter localization was therefore performed using the `u-track` software package originally developed for single-particle tracking[176]. To remove false positive localizations from results prior to downstream analysis, detected emitters were filtered based on their localization properties.

Channel registration: Images from different spectral channels have to be aligned since aberrations caused by optical components of the microscope distort light with different wavelengths to a variable degree. This can be done by transforming raw images or based on coordinates of objects detected in both spectral channels. In the case of the gSEP construct, the presence of the reference marker and the protein tags within one molecule provide signals in both spectral channels which can be directly used for channel registration. Registration was performed directly on localizations obtained from emitter detection by computing translation, rotation and scaling factors to minimize the distance between signals in the reference and target channels (Fig. 3.3e).

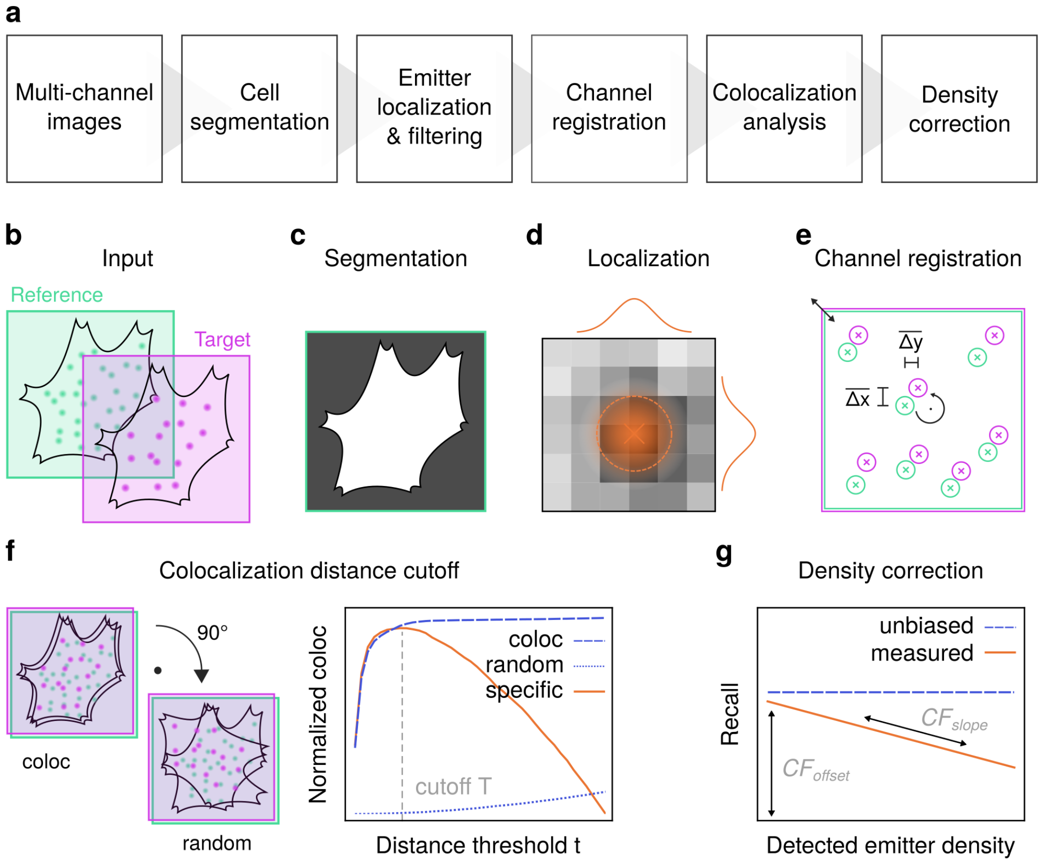

Figure 3.3 Data processing steps for single-molecule labeling efficiency measurements. a, Data processing workflow. **b,** Raw multi-channel images for reference and target channel are used as input for cell segmentation and emitter localization. **c,** Cell segmentation based on reference channel. **d,** Gaussian fitting of diffraction-limited signals gives access to sub-pixel location, signal amplitude and width. **e,** Channel registration scaling, rotation and translation using sub-pixel coordinates of reference and target localizations as input. **f,** Computation of colocalization distance cutooff. Left: Target channel coordinates are rotated by 90° to generate randomized coordinates. Right: Normalized distance-dependent colocalization between reference and target channels. **g,** Correction of density-dependent emitter detection performance.

Colocalization analysis: Prior to computing the fraction of colocalizing emitters in the reference vs. target channel, a distance cutoff is required to distinguish between specific colocalization of signals from the same gSEP molecule and random colocalization of signals originating from different gSEP molecules. This cutoff depends on the registration quality between reference and target channel, as well as on the localization precision for individual emitters in both channels. At small cut off values below the localization precision, even small localization errors would cause a large fraction of true co-localizations to not be identified. At high cutoff values, reference and target localizations from different gSEP molecules may be falsely assigned as colocalizing signals. The specific cutoff distance thus maximizes the number of true colocalizations while minimizing the number of random colocalizations.

The specific cutoff distance $coloc_{specific}(t)$ was determined by computing the difference between the number of colocalization events, $coloc = localization_{ref} \cap localization_{target}$, observed in datasets after registration of reference and target channel and in a randomized dataset across different distances ranging from 0 to t_{max}:

$$coloc_{specific}(t) = \frac{coloc(t)}{coloc(t_{max})} - \frac{coloc_{random}(t)}{coloc_{random}(t_{max})}. \tag{3.2}$$

Here, the randomized dataset where no specific colocalization was expected was generated by rotating coordinates of localizations in the target channel by 90° (Fig. 3.3d). Fig. 3.3f shows the typical distance-dependent behavior for $coloc(t)$, which shows a steep increase at small distances t and for $coloc_{random}$ which shows a more gradual increase with increasing distance t.

The observed maximum of $coloc_{specific}$ at intermediate t was used as the distance cutoff T for which reference and target channel localizations were considered to colocalize.

Density correction: In principle, the labeling efficiency can be computed based on the observed degree of colocalization between reference and target channel emitters located closer than T to each other and the total number of detected emitter in the reference channel (Equation 3.1). This approach will report unbiased labeling efficiency measurements if all emitters in the reference and target channel are successfully detected at their true positions. However, diffraction limits our ability to directly determine the position of individual emitters and additional confounding factors such as overlapping emitter signals, a limited SNR or variable background features will occur in experiments. It is therefore typically not possible to localize all emitters without introducing positional errors and without missing a fraction of emitters altogether. This means that in order to achieve unbiased labeling efficiency measurements by single-molecule colocalization analysis, measurements have to be adjusted for imperfect emitter localization.

In simulations with varying emitter density, we previously observed that the recall, that is the fraction of simulated emitters that were successfully localized, was anti-correlated with the simulated emitter density and consequently also with the detected emitter density (Fig. 3.3g). Here, the anti-correlation between recall and detected emitter density exhibited a linear dependence. As the detected emitter density is readily available after emitter localization, this means that the relation between detected emitter density and recall can be used to correct for density dependent errors in emitter localization.

To enable the correction of the degree of colocalization determined from experimental data, images containing simulated emitters with properties matching experimental data were simulated for different emitter densities. These simulated images were then subjected to emitter localization and the relation between recovered emitter density and simulated emitter density was linearly fitted to extract the slope and offset as correction factors.

The Degree of colocalization (DOC) computed from individual images was then corrected to compute the Labeling efficiency (LE) according to:

$$LE = 100\frac{DOC}{CF_{slope}dens_{meas} + CF_{offset}} \tag{3.3}$$

where CF_{slope} and CF_{offset} are the correction factors obtained from the analysis of simulated images and $dens_{meas}$ is the detected emitter density in the target channel of the analyzed image.

3.3 Scope

The goal of this study was to develop and validate a method for measuring absolute labeling efficiencies in the context of protein counting with single-molecule fluorescence microscopy. To achieve this, I built on previous work performed by Siegfried Haenselmann in our lab, who established gSEP as validation tool for single-particle tracking and for identifying optimal labeling conditions for multicolor single-particle tracking[174]. With an initial version of the data analysis pipeline described above, maximum labeling efficiencies of 40 % for $\mathrm{SNAP_f}$-tag and 50 % for HaloTag were measured. It remained unclear whether this limit in labeling efficiency was due to inherent limitations of the applied data analysis routines, or if the labeling efficiency achieved with protein tags depends on the tags, fluorescent ligands or biological systems in which they are being used.

To understand how technical limitations of the single-molecule colocalization routine influence measured labeling efficiencies, the first aim was to validate and improve the data analysis pipeline. This included comparing the current emitter localization approach with alternative approaches, and validating the analysis routine with simulations and experimental data. Once a validated analysis pipeline was established, the next aim was to automate data acquisition to facilitate systematic studies of protein tag performance across different labeling conditions. The last aim of the work presented here was to study the influence of fluorescent ligands, fixation conditions and host cell lines on the performance of $\mathrm{SNAP_f}$-tag and HaloTag. The observations made in this context will guide the choice of labeling conditions for in-situ protein counting with photobleaching step analysis in chapter 4.

3.4 Optimized labeling efficiency measurements with gSEP

In previous work with gSEP, image sets were acquired manually for each individual cell by single-molecule sensitive TIRF microscopy. Manual acquisition limits the amount of cells that can be imaged per sample and in addition may result in photobleaching during focusing and cell selection. As part of this study, I set out to automate data acquisition by establishing the `ImageJ`-based microscope control software `µManager`[177]. To enable fully unsupervised image acquisition, the microscope was additionally equipped with a motorized emission filter wheel. This allowed for scripted acquisition of multi-channel images according to the original gSEP imaging protocol for >20 cells per condition across up to eight conditions in one sample holder. If required, the same imaging routine could be extended towards additional conditions if larger sample holders are used. One important deviation in the acquisition routine was that cells for imaging were pre-selected prior to data acquisition based on the `Slide Explorer` plugin of `µManager`. This reduced the light exposure of the sample prior to data acquisition to 50 ms per cell and thereby substantially decreased the risk of photobleaching. Using the automated acquisition script, the typical duration for an experiment with eight conditions was reduced from about one day of work to about one hour of preparation time for cell selection and about six hours of unsupervised imaging time while increasing the number of imaged cells per condition by at least 2-fold.

For data processing, `DOLanalysis` was developed as object-oriented analysis framework in MATLAB. This framework is centered around emitter localization sets ('pointsets'), multichannel images and sets of multichannel images ('imagesets') as core objects and was used for all data processing performed in this chapter. The principles behind each processing step remained unchanged from the initial version of the gSEP approach described in Section 3.2.2. Specific procedures for individual experiments and critical parameters during analysis are described in Section 6.3.2. Taken together, the established approach for semi-automated data acquisition and flexible analysis framework provided the basis for optimizing and validating individual steps during analysis as will be described below.

3.4.1 Improved emitter localization with thunderSTORM

Detection of diffraction-limited signals and sub-pixel localization are the basis for measuring labeling efficiencies with gSEP since they convert raw images into localizations that are subsequently used for all downstream analysis steps. It is therefore critical that this emitter detection step of the data analysis routine produces accurate results, which are as robust as possible towards changes in input data and exhibit as little as possible variation across cells. In the initial version of the gSEP approach, emitter detection was performed using the sub-pixel localization routine of the `u-track` package for single-particle tracking[176]. It was observed that while this emitter detection and localization algorithm generally performed well, it also returned putative emitter positions which did not correspond to diffraction-limited signals identified by eye. This subset of detection events could be largely removed by applying an intensity filter with empirically determined threshold, but their origin as well as the fraction of low-intensity localizations under varying experimental conditions was difficult to predict.

For this reason, I decided to cross-validate emitter detection obtained with `u-track` with emitter localizations obtained with an alternative localization framework. After comparing the performance of different localization frameworks on representative experimental data recorded on different gSEP datasets, I chose the `thunderSTORM` software package[178] for a detailed comparison with `u-track` emitter localization. `thunderSTORM` is available as a plugin for `ImageJ` and therefore easy to adapt for automated processing of large datasets. `thunderSTORM` additionally allows to perform multi-emitter fitting where signals from partially overlapping emitters are modeled with a mixture of multiple Gaussians to enable correct localization of each individual emitter and to avoid misplaced localizations. Emitter detection settings for `thunderSTORM` were optimized by visual inspection and kept fixed for all further analysis. Specific detection parameters are given in section 6.3.2.

3.4.2 Benchmarking of emitter localization frameworks with simulations

To achieve a systematic evaluation of the performance of `thunderSTORM` and `u-track` across different emitter densities and background strengths, I compared both frameworks with the help of simulated data. For simulated data, the ground truth positions of emitters are known and a comparison of detected emitter locations with the ground truth can be performed. Unlike an evaluation with experimental data, evaluation with simulated data enables a quantitative comparison of emitter localization approaches. One prerequisite for this is that the simulated data used for benchmarking closely resemble the experimental data which will be processed with the established analysis routine.

Simulated images resembling cells labeled with protein tag ligands, were obtained by combining real images of cells expressing gSEP without labeling of protein tags with simulated images containing diffraction-limited signals (Fig. 3.4a).

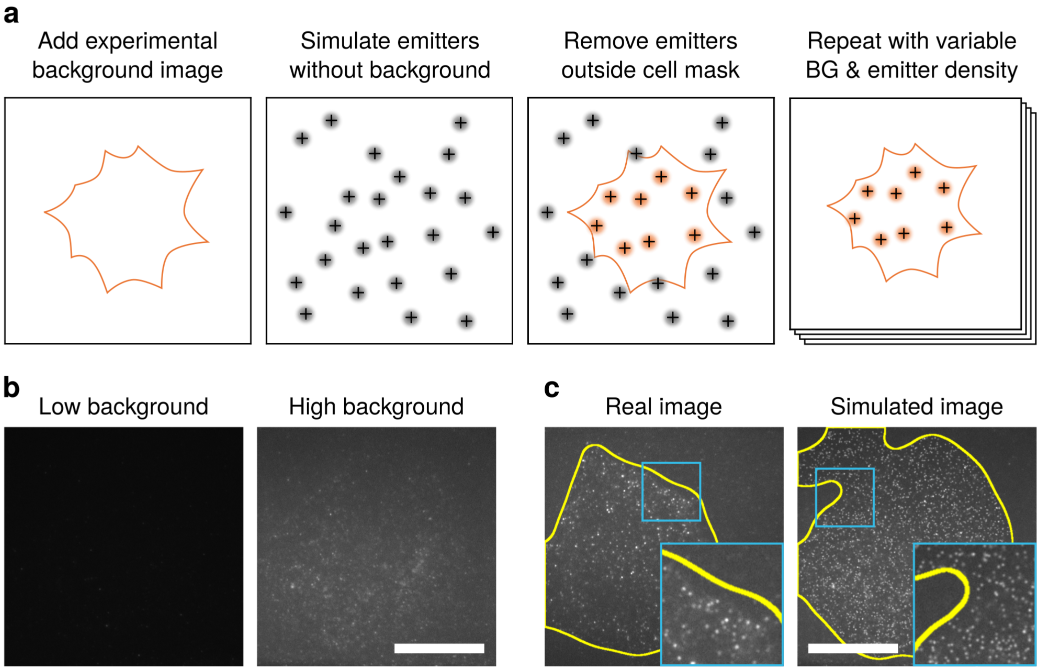

Figure 3.4 Routine for generating simulated input data for emitter detection benchmarking. a, Overview data generation scheme. **b,** Representative input data from high background dataset. Left: low background image acquired on unstained H838 cell expressing gSEP construct under TIRF illumination and at 561 nm with acquisition settings identical to those used for measurements on labeled cells. Right: High background image of same cells as for low background with 640 nm excitation settings. **c,** Comparison of image from real cell (left) and simulated cell (right) from high background dataset with overlayed cell segmentation (yellow lines). Scale bars: 20 µm. Simulations were established jointly with Felix Braun.

To emulate background signals typically observed in fluorescence microscopy images of mammalian cells, we compiled a set of representative images of unlabeled cells from different gSEP cell lines acquired under identical acquisition settings as were used for imaging gSEP target channels. Diffraction-limited signals mimicking the signals originating from individual labeled gSEP molecules in cells after labeling of protein tags were generated using the software `testSTORM`[179]. `testSTORM` enables simulations of individual emitter signals with adjustable signal strength, photophysics and size and is therefore ideally suited to generate test data which closely resembled our experimental data (see Section 6.3.3 for simulations parameters). Using `testSTORM`, we simulated emitter signals randomly distributed across images at varying densities between 0 and 2.6 μm^{-2}. Simulated emitters located outside the area covered by cells were removed by applying a mask, which was generated based on the gSEP reference (EGFP) channel.

The background levels observed in the two target channels with 561 nm and 640 nm excitation appeared different in overall signal levels, as well as in the number of distinct features typically contained in cells. We therefore simulated two datasets with low and high background, where the low background scenario corresponded to background signals typically encountered in the 561 nm excitation target channel, while the high background scenario corresponded to background observed in the 640 nm excitation target channel (Fig. 3.4b). Images resulting from this simulation approach appeared qualitatively comparable to real data recorded on cells in which gSEP had been labeled with respective ligands for $\mathrm{SNAP_f}$-tag and HaloTag (Fig. 3.4c).

Simulated images with low ($0.2\,\mu\mathrm{m}^{-2}$) or high ($1.6\,\mu\mathrm{m}^{-2}$) simulated emitter densities were then subjected to emitter detection by `u-track` and `thunderSTORM` using detection parameters previously determined for experimental data. Overall, I observed a good performance with both algorithms at simulated emitter densities $\leq 0.5\,\mu\mathrm{m}^{-2}$ (Fig. 3.5a), especially in the low background scenario. As previously observed in experimental data, `u-track` showed a tendency to detect emitters at positions with minimally increased background where no simulated emitters were located (Fig. 3.5a, arrow 1). A small fraction of dim simulated emitters was missed by both frameworks (Fig. 3.5a, arrow 2). Bright features in simulated images, which originated from used experimental background images and not from simulated emitters were detected by both frameworks (Fig. 3.5a, arrow 3). Since these feature appeared to be highly similar in size and brightness to simulated emitters, it is likely that a certain degree of localizations from autofluorescent particles will also be encountered in experimental data.

At higher simulated emitter densities, both detection frameworks failed to identify simulated emitters to a varying degree (Fig. 3.5b). Multi-emitter fitting with `thunderSTORM` helped to correctly identify simulated emitters located in close proximity. Such emitters were typically missed by `u-track`, which tended to place localization between closely located simulated emitters (Fig. 3.5b, arrow 4). If simulated emitters overlapped strongly or if a large number of simulated emitters occurred in close proximity, both detection frameworks typically started to miss individual emitters and `u-track` increasingly failed to place any localizations at all in these areas (Fig. 3.5b, arrow 5). In contrast to the performance of `u-track` at low simulated emitter densities, the high sensitivity of `u-track` seemed to be

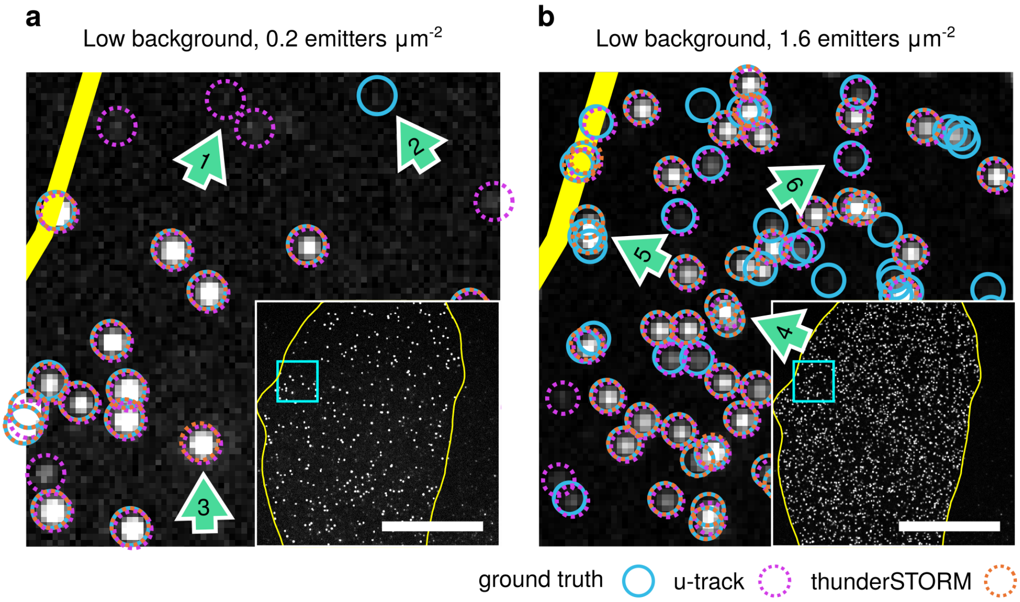

Figure 3.5 Qualitative comparison of emitter detection performance. Emitter detection with `u-track` and `thunderSTORM` localization algorithms at low (a) and high (b) simulated emitter densities. Zoomed regions of full image shown in inset. Emitter positions were omitted from inset. Yellow lines indicate masked region of cell. Arrows indicate notable observations in dataset. 1: Detection of background by `u-track`. 2: ground truth emitter missed by both algorithms. 3: bright autofluorescent spot localized by both algorithms. 4: Two closely located emitters correctly detected by `thunderSTORM` and missed by `u-track`. 5: Group of closely located ground truth emitters not correctly localized by both algorithms. 6: Dim ground truth emitter only localized by `u-track`. Scale bars: 20 µm.

beneficial at higher emitter densities since `u-track` tended to identify low intensity emitters at a higher success rate than `thunderSTORM` (Fig. 3.5b, arrow 6).

Overall, the qualitative evaluation of `u-track` and `thunderSTORM` showed that under good conditions (low emitter densities and low background) both algorithms succeeded in identifying the majority of simulated emitters. However, upon increasing simulated emitter densities and background, `thunderSTORM` seemed to perform better. To which degree the high sensitivity of `u-track` negatively impacts the results of labeling efficiency measurements remained difficult to assess. This suggests that, for experimental conditions with low background signal and low-intermediate emitter densities, emitter localization with `thunderSTORM` represents an improvement over emitter localization with `u-track`.

Simulated data allow for a quantitative evaluation by comparison to the known position of simulated emitters. I therefore proceeded to systematically compare both detection frameworks across simulated emitter densities between 0.1 and $2.6\,\mu\mathrm{m}^{-2}$ under the low and high background scenario. For emitter localization in the context of labeling efficiency measurements two metrics are critical: The recall and the false positive density. The recall refers to the fraction of simulated emitters correctly identified. It is the primary criterion since since efficient detection of emitters is crucial for robustly detecting colocalization events. The recall is mainly driven by the SNR of emitter signals. The false positive density refers to the number of false-positive localizations per area identified in images. In contrast to the recall, the false positive density tends to depend more strongly on background signals since these can be difficult to distinguish from simulated emitters (Fig. 3.5).

Since it is difficult to compare the recall and false-positive density against each other, the Jaccard Index (JI) can be used as a high-level metric for assessing the detection performance of localization algorithms[175]. The Jaccard index is defined as:

$$JaccardIndex(JI) = \frac{TP}{TP+FP+FN} \tag{3.4}$$

where TP is the number of true positive localizations, FP is the number of false positive localizations and FN is the number of false negative localizations. To determine FP, I performed a colocalization analysis between the ground truth emitter positions (reference channel) and recovered emitter positions (target channel) as described in Section 3.2.2. Localized emitters detected as colocalization events were used as TP, while ground truth emitters without colocalizing target emitters were considered to be FN. Surplus localized emitters not colocalizing with ground truth emitters were used as a measure of FP.

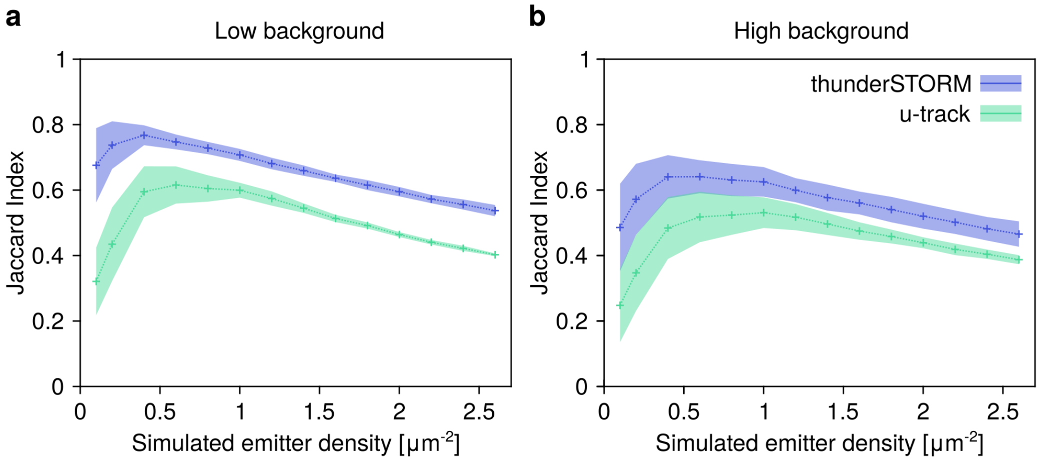

Figure 3.6 Quantitative comparison of emitter detection with utrack and thunderSTORM. Jaccard indices for emitter localization with `u-track` or `thunderSTORM` at variable emitter densities from simulations with low (**a**) and high (**b**) background strength. Shaded regions indicate standard deviation around means (crosses) at indicated emitter densities. For each algorithm, 20 simulated images per condition were analyzed.

In general, JI for emitter localization with `thunderSTORM` were found to be ≥10 % higher across all simulated emitter densities and in both background scenarios (Fig. 3.6). Both emitter localization algorithms exhibited a maximum in their emitter detection performance at intermediate simulated emitter densities (∼0.5-1.0 μm^{-2}). With a peak JI of 0.77 ± 0.03 at a simulated emitter density of 0.4 μm^{-2}, `thunderSTORM` tended to perform better at slightly lower emitter densities, while `u-track` exhibited its peak performance at simulated emitter densities of 0.6-1.0 μm^{-2}. At simulated emitter densities ≤0.4 μm^{-2}, `u-track` showed a strongly decreased performance (<50 % peak JI), while `thunderSTORM` showed ∼10 % lower peak JIs at simulated emitter densities of 0.2 μm^{-2}. At simulated emitter densities >1 μm^{-2}, both algorithms showed a linear decrease in achieved JI in both background scenarios. The overall performance of both algorithms was worse at higher background levels, while the influence of background strength on JI was lower for `u-track` than for `thunderSTORM`.

To assess the causes for the observed emitter density-dependent variations in JI, I additionally compared the recall and false-positive density across simulated emitter densities and background scenarios (Fig. A.1). This analysis revealed that the recall for both localization algorithms exhibited an almost linear dependence on simulated emitter densities for both background scenarios. In line with the JI, higher recalls were achieved with `thunderSTORM` compared to `u-track` and both algorithms achieved higher recalls in the low background scenario. As already observed in exemplary data shown above, the false-positive density was high for emitter localization with `u-track` at low emitter densities explaining the pronounced decrease in JI at low simulated emitter densities. The same tendency to produce false-positive localizations at low simulated emitter densities was observed for `thunderSTORM`. Since the overall false-positive detection rate for `thunderSTORM` was at least 2-5-fold lower, this behavior was not noticed during manual inspection of localization performance. Interestingly, the false-positive density observed for `u-track` at low simulated emitter densities exhibited a strong image-to-image variability, suggesting that some background images contained in the dataset showed fewer features causing false-positive `u-track` localizations. At increasing simulated emitter densities, the false-positive density increased steadily for both localization algorithms and in both simulated background scenarios.

To conclude, replacing the previously used emitter detection framework `u-track` with `thunderSTORM` resulted in an overall improvement in emitter detection and also in an improvement in density-dependent performance of emitter localization on simulated data.

3.4.3 Localization filtering

Although replacing the emitter detection framework led to improved detection efficiencies, a population of false-positive localizations was observed when processing simulated data where emitters had a well-defined size and shape. In experimental data, additional sources of undesired signals which were not included in our simulations, such as emitters located outside the focal plane or emitter aggregates are likely to cause additional false-positive localizations.

Sub-pixel localization of emitters via fitting of 2D-Gaussians gives access to additional information about each emitter. In SMLM localization properties such as the signal amplitude, the localization precision or the width of the fitted Gaussian are routinely used to improve the quality of reconstructed, super-resolved images. In the previous version of the gSEP data analysis pipeline, an amplitude filter was used to remove false-positive localizations. However, the intensity under TIRF illumination strongly depends on the distance of an emitter from the coverslip surface and additional illumination intensity variations due to a spatially variant illumination profile are frequently observed when performing TIRF microscopy on cellular samples (see e.g. Fig. 4.12). In addition, I observed that localization amplitudes varied between experiments and sometimes even within individual datasets (see e.g. Fig. A.4c).

A second parameter that is readily obtained from sub-pixel localization is the width (sigma) of individual signals. In contrast to the amplitude, sigma is mainly governed by the optical components in a microscope and closely related to the diffraction limit of the optical system (section 2.1). If the majority of signals in an image is limited by diffraction, values smaller or larger than the typical sigma value indicate that a signal does not originate from an individual, diffraction-limited emitter. I therefore tested, if a filter to remove localizations with deviating sigma could be used to remove false-positive localizations, including localizations of background noise, out-of-focus emitters and emitter aggregates (Fig. 3.7a).

To investigate how filtering emitter localizations based on their fitted sigma values impacted downstream colocalization analysis, I first examined the distribution of sigma values observed in experimental data upon emitter localization with `thunderSTORM`. Fig. 3.7b,c show representative sigma distributions for localizations in the reference and target channel of a dataset recorded in H838 gSEP cells. Sigma values in both channels typically exhibited unimodal distributions with a skew towards larger sigma values. In line with expectations based on the used excitation wavelengths of EGFP (488 nm) and BG-SiR (640 nm), the mode of the sigma distribution for BG-SiR was shifted towards higher sigmas (1.225 pixel, Fig. 3.7c) compared to the sigma distribution observed for EGFP (mode: 1.025 pixel, Fig. 3.7b). In both channels, a small fraction of localizations with sigma values ≤ 0.1 pixel was observed, which presumably represent localizations of background noise. A second fraction of putative false-positive localizations was characterized by sigma values substantially larger than the mode of the sigma distribution. These localizations potentially could have been caused by signals which contained multiple emitters in close proximity as they were for example observed during analysis of simulated images with high simulated emitter densities (Fig. 3.5b). Although it was clear that both, localizations with low and high sigma values could lead to biased labeling efficiency measurements, precise cutoff values which distinguished between false-positive and true positive localizations could not be identified based on the sigma distribution alone.

I therefore proceeded by applying filters centered around the mode of the sigma distribution for all emitters in a given detection channel. I further reasoned, that since undesired localizations should occur equally in both, the reference and target channels, equal relative filter widths had to be applied to both channels. Upon increasing the width of the applied sigma filter window as fraction of the mode (mode$\pm$X%), the fraction of included emitter localizations in both channels increased (Fig. 3.7d). In conjunction with increasing fractions of emitters included after filtering, the computed degree of colocalization between emitters in

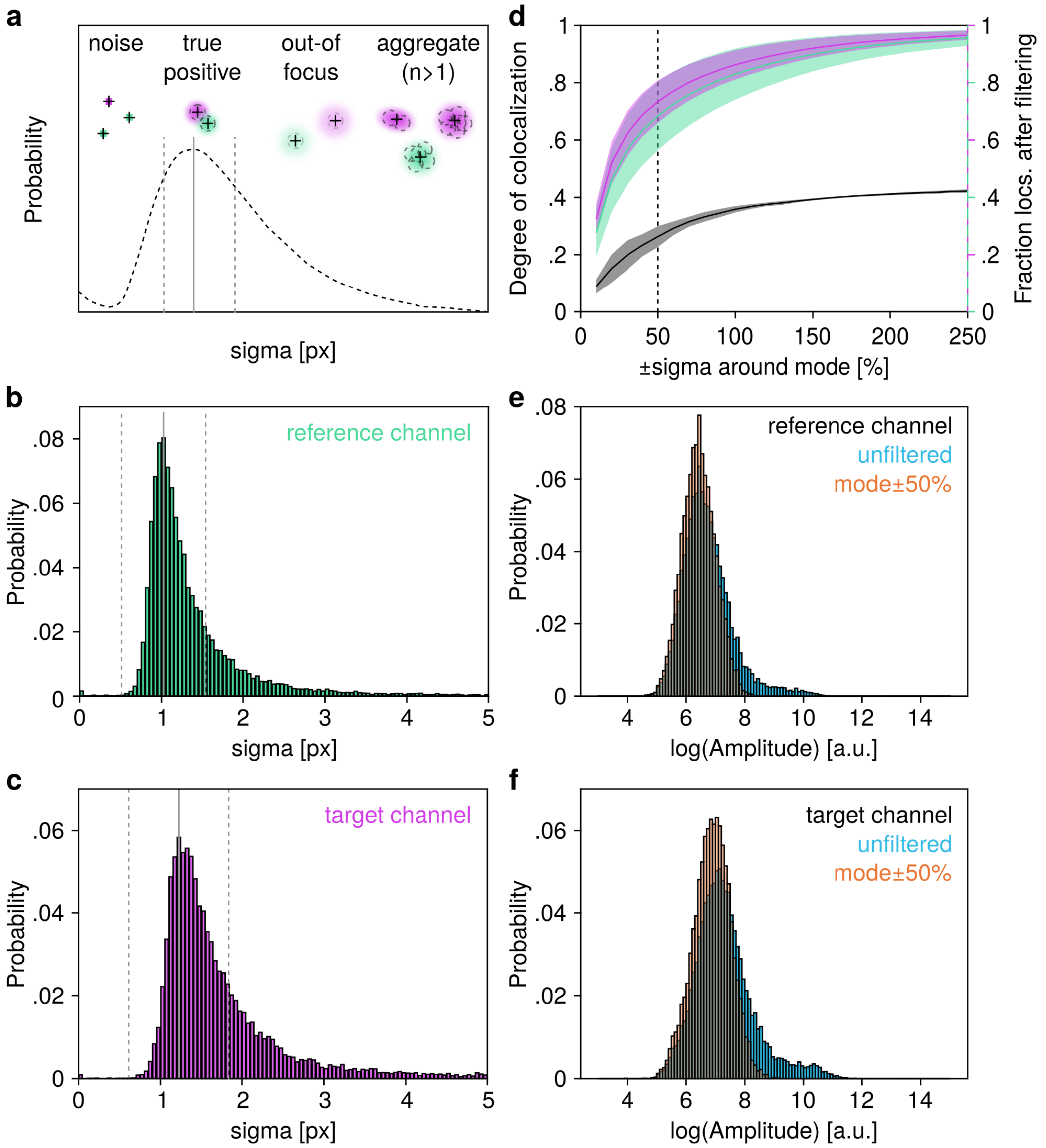

Figure 3.7 Post-detection localization filtering based on emitter signal width. **a,** Schematic overview of signal sources and expected sigma values across typical sigma distribution. Crosses: centroid, dashed circles: sigma. **b,c** Sigma distribution in reference (b) and target (c) channel obtained from H838 gSEP cells labeled with BG-SiR (500 nM) for 3 h. Gray line indicates largest bin around which applied filter is centered. Dashed lines indicate width of filter for window size of ±50 %. **d,** Degree of colocalization between reference and target channel localizations upon sigma filtering and fraction of retained localizations after filtering with varying window size for reference (green) and target (magenta) channel. The degree of colocalization was computed at a constant distance cutoff T of 3.0 pixel. **e,f** Amplitude distributions for localizations in reference (e) and target (f) channel before (blue) and after (orange) applying sigma filter with ±50 % window width.

the reference and target channel also increased. As this analysis showed a gradual dependence of the degree of colocalization on the applied filter width, it was not possible to establish cutoff value separating false positive from true positive localizations.

Although the typical amplitude of emitters in a dataset varies depending on a number of variables (see above), the distribution of amplitudes in single-molecule fluorescence microscopy experiments typically follows a log-normal distribution[106]. When assessing the amplitude distributions for emitter localization in the reference or target channel of measurements with gSEP, I observed that amplitude distributions typically contained a large population following a log-normal distribution, as well as a second smaller distribution with high amplitudes (Fig. 3.7e,f). Interestingly, this second population was largely removed by applying a filter in the sigma channel with a filter window width of $\leq \pm 50\,\%$ around the mode (orange distributions in Fig. 3.7e,f). As this behavior could be reproduced across different datasets, I decided to use a sigma filter with a filter width of $\pm 50\,\%$ (indicated as dashed lines in Fig. 3.7b,c) for further evaluations with simulated data.

To evaluate the density-dependent influence of the established sigma filtering approach on the overall emitter localization performance, I reanalyzed the simulated image sets I initially used for comparing emitter localization with `u-track` and `thunderSTORM` (Section 3.4.2). For the reanalysis with `thunderSTORM`, I applied the mode$\pm 50\,\%$ sigma filter described above. For reanalysis with `u-track`, I combined previously established amplitude filtering[174] with the sigma filtering approach as was used for filtering of `thunderSTORM` localizations. As previously, I then compared the performance of both localization frameworks after filtering based on their recall, false-positive density and the Jaccard index to study the impact of filtering.

Applying additional filters post emitter detection substantially changed the performance of emitter localization with `u-track`. The achieved Jaccard index for `u-track` with filtering at low simulated emitter densities was almost doubled in both, the low and the high background scenario (Fig. 3.8). This was mainly due to a substantially reduced false-positive detection density (Fig. A.1c,d). For `thunderSTORM` applying an additional post localization sigma filter had a negligible effect in the low background scenario across all simulated emitter densities ($\leq 4\,\%$ difference). Under the high background scenario, applying a post localization sigma filter improved the Jaccard index by 5-8 % at simulated emitter densities $<0.4\,\mu\mathrm{m}^{-2}$. Among the tested combinations of emitter localization and filter approaches, `thunderSTORM` in combination with the established sigma filtering approach exhibited the best overall performance with respect to all three evaluated metrices and across the large majority of

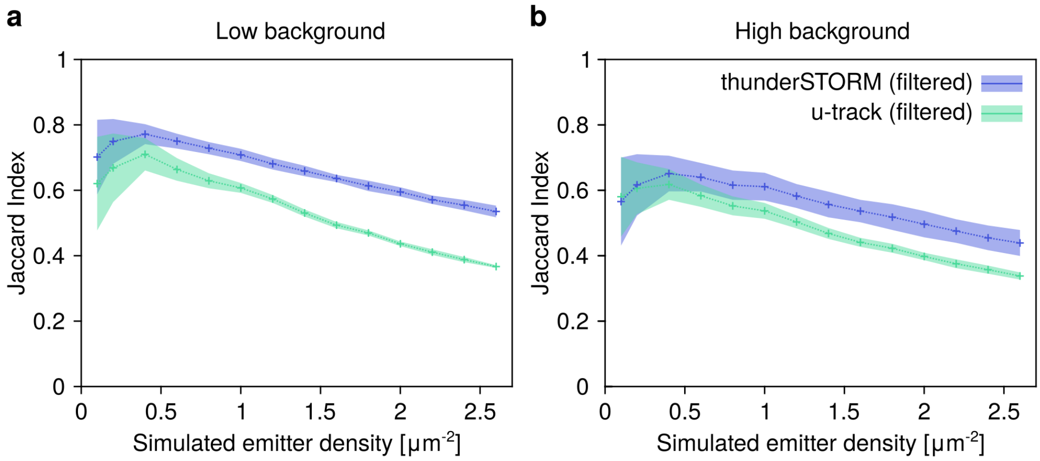

Figure 3.8 Impact of localization filtering on emitter detection. Jaccard indices after localization filtering at variable emitter densities from simulations with low (**a**) and high (**b**) background strength. Shaded regions indicate standard deviation around means (crosses) at indicated emitter densities. For each algorithm, 20 simulated images per condition were analyzed.

simulated emitter densities. I therefore decided to use `thunderSTORM` in combination with the applied sigma filtering approach (mode±50 %) as emitter localization framework in all further experiments presented in this study.

3.5 Validation of the optimized DOLanalysis pipeline

After having established that replacing the initial `u-track` emitter localization approach with `thunderSTORM` and applying a localization filtering step improved the emitter detection efficiency and specificity in simulations, I set out to benchmark the updated `DOLanalysis` pipeline with three complementary approaches.

3.5.1 Reference emitter densities in wild-type cells

In the first evaluation experiment, I compared the specificity of emitter localization in the EGFP reference channel after emitter localization with `thunderSTORM` and sigma-based localization filtering. For this, I prepared and imaged Huh7.5 cells expressing gSEP or LynG with Huh7.5 wt cells where no specific EGFP signal was expected.

The detected emitter density in both, Huh7.5 gSEP and LynG cells varied considerably with localization densities after localization filtering ranging from ~0.01 to 1.1 μm^{-2}.

The median detected reference emitter density in Huh7.5 gSEP cells was found to be $0.75 \pm 0.33\,\mu\text{m}^{-2}$, while Huh7.5 LynG exhibited a median emitter density of $0.40 \pm 0.30\,\mu\text{m}^{-2}$. This means that in a typical Huh7.5 gSEP cell more than 1×10^3 individual gSEP constructs were detected, while in a typical Huh7.5 LynG cell, around 6×10^2 LynG constructs were detected. In contrast, Huh7.5 wild-type cells exhibited a median detected emitter localization density of $0.013 \pm 0.008\,\mu\text{m}^{-2}$ (or ~20 localized emitters per cell) and showed less cell-to-cell variation. As the background signal was previously observed to be worse in target channels with excitation at 640 nm, I additionally assessed the detected emitter density at 640 nm excitation in wild-type cells. Also here, a very low detected emitter density of $0.008 \pm 0.014\,\mu\text{m}^{-2}$ was observed.

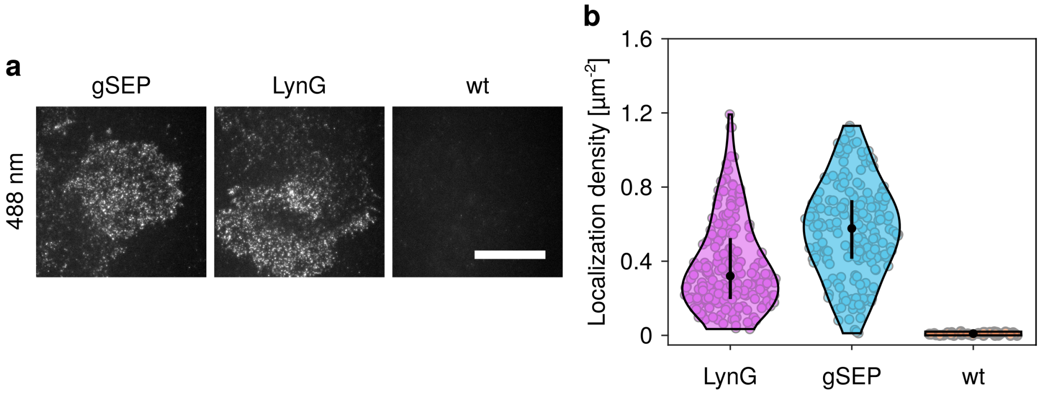

Figure 3.9 Comparison of detected emitter densities in reference channel after localization filtering. Huh7.5-gSEP, -LynG and wild-type (wt) cells were prepared and imaged with standard settings (Section 6.2.3) and sigma-based filtering of localizations was performed. **a,** Representative images from indicated cell types. All images acquired and displayed with identical settings. Scale bar: 20 µm. **b,** Detected emitter densities from indicated cell types shown as total number of localization detected per cells divided by the segmented area of the respective cell. Circles represent individual cells. Data from 194 (gSEP), 189 (LynG) or 38 (wt) cells from 5,6 or 1 samples.

This analysis first of all shows that emitter localization with `thunderSTORM` in combination with localization filtering based on $\sigma$was able to achieve very high specificity. On average, more than 96 % (98 %) of localizations in Huh7.5 gSEP (LynG) cells originating from expressed gSEP (LynG) constructs. Secondly, this analysis demonstrates that for the selected imaging and emitter localization conditions, very low false positive detection rates, caused by autofluorescence or background signal in the reference channel, are to be expected.

3.5.2 Dynamic range of DOLanalysis

For the second validation step I again used simulated data to assess the dynamic range of the `DOLanalysis` pipeline. As before (Section 3.4.2), images of cells containing diffraction-limited signals were generated to emulate signals from gSEP constructs after protein tag labeling with fluorescent ligands. In contrast to previous simulations, this time emitters were selectively removed from images to generate a set of target channel images with variable emitter numbers. By providing the ground truth locations as reference localizations, image set of cells exhibiting varying labeling efficiencies ranging from 5-95 % were obtained. Following my previous observations that different background levels occurred in the target channels upon excitation at 561 nm and 640 nm (Section 3.4.2), a set of images for the low and high background scenario was simulated. In addition, the simulations were repeated at two different ground truth emitter densities of 0.6 and 1.6 μm^{-2}. A reference channel emitter density of 0.6 μm^{-2} was typically encountered across the different cell lines used in this study, while a density of 1.6 μm^{-2} emulated the situation in cells with strong gSEP expression.

The generated image sets were then analyzed by performing emitter localization and filtering with `thunderSTORM` followed by channel registration, colocalization analysis and density correction using `DOLanalysis` as described above. The recovered median labeling efficiencies for each set of images were then plotted against the simulated labeling efficiency for each of the four combinations of background scenario and reference emitter density Fig. 3.10.

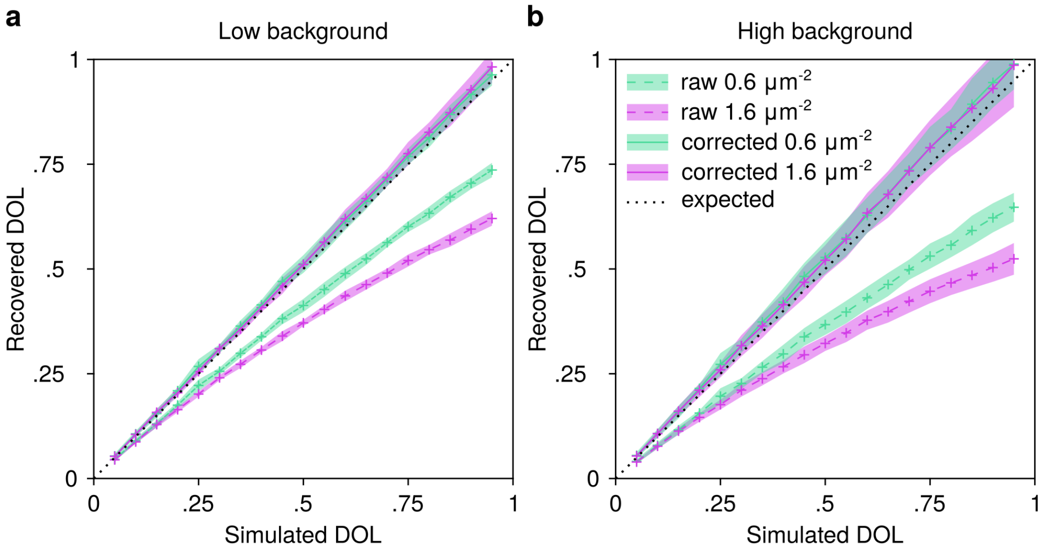

Figure 3.10 Validation of gSEP pipeline with simulations. Images with specific DOL were generated by randomly selecting localization subsets and generating cell images from these subsets with low (a) or high background (b) at a ground truth emitter density of 0.6 or 1.6 μm^{-2} as described in Section 3.4.2. Generated images were analyzed with the `DOLanalysis` pipeline as described above and are shown as median (crosses) ± SD (shaded region) before (raw) and after (corrected) density correction. Black dashed line: expected slope for error-free recovery of simulated DOL values. 20 images were simulated and analyzed per condition. Data was produced jointly with Felix Braun.

As expected, labeling efficiency estimates prior to density correction systematically deviated from simulated labeling efficiencies. The degree of mismatch between simulated and recovered labeling efficiency varied depending on overall emitter density and the chosen background scenario. As expected, the mismatch between recovered and simulated labeling efficiencies increased with higher simulated labeling efficiency. This is expected, as both, the simulated ground truth density as well as the labeling efficiency influence the number of emitters occurring in the reference channel images. Consequently, increasing emitter numbers will result in decreased overall emitter localization efficiency. On average, the recovered labeling efficiencies were on average ∼10 % higher in the low background scenario compared to the high background scenario, which is in line with my previous observation that `u-track` performed better in the low back ground scenario. After density correction, the recovered labeling efficiencies closely matched the simulated labeling efficiencies for all four simulated scenarios. A small, yet systematic tendency towards over-correcting recovered labeling efficiencies was observed. However, the average deviations between simulated and recovered labeling efficiencies in this case were below 4 % in all four scenarios and across all simulated labeling efficiencies. Interestingly, the image-to-image variability was higher for datasets from the high background scenario, which indicates that the performance of emitter localization with `thunderSTORM` and sigma-filtering may also depend on specific features of individual cells.

Overall, this analysis confirmed that despite the emitter density-dependent performance of the employed emitter localization algorithms the analysis pipeline works as intended. I therefore conclude that labeling efficiencies covering at least the range between 5 and 95 % could be successfully measured using the established data analysis pipeline.

3.5.3 Experimental validation with multispectral beads

Having confirmed that the chosen imaging approach allowed for acquisition of emitter signals from gSEP constructs with high specificity and that labeling efficiencies covering the range from 5-95 % could be recovered with `DOLanalysis`, I sought to show that `DOLanalysis` was able to measure high labeling efficiencies approaching 100 % in experimental data.

As samples for this experiment, I chose multispectral fluorescent beads with sub-diffraction size (100 nm) which contain different fluorophore species that can be excited at 488 nm, 561 nm and 640 nm, i.e. in the reference and both target channels used in this study. They also contain multiple fluorophore molecules for each fluorophore type, which makes them suited validation targets as the degree of colocalization in this case is expected to approach 1. To avoid confounding effects from the applied density correction routine, I performed this analysis with samples prepared to contain a low density of beads randomly deposited on the surface of coverslips. Excitation intensities were chosen to result in emitter amplitudes after localization which were comparable to amplitudes typically observed in cell experiments with gSEP.

Figure 3.11a shows a representative overlay of reference and emitter channel images. Typically, a high degree of overlap between both images was observed in the center of images, while aberrations resulted in less overlay towards the edges of the Field of view (FOV). The median detected emitter density in the reference channel was $0.26 \pm 0.01\,\mu\text{m}^{-2}$ across all images contained in the analysis. As expected, the majority of observed signals in both, the reference and target channel appeared to be symmetric and well-separated (Fig. 3.11b,c). Consequently, a large fraction of overlap between emitter localizations in both channels was obtained after channel registration (Fig. 3.11d). This was also reflected by the distribution of nearest neighbor distances obtained from colocalization analysis, where more than 90 % of emitter localizations in the reference channel after channel registration exhibited a distance

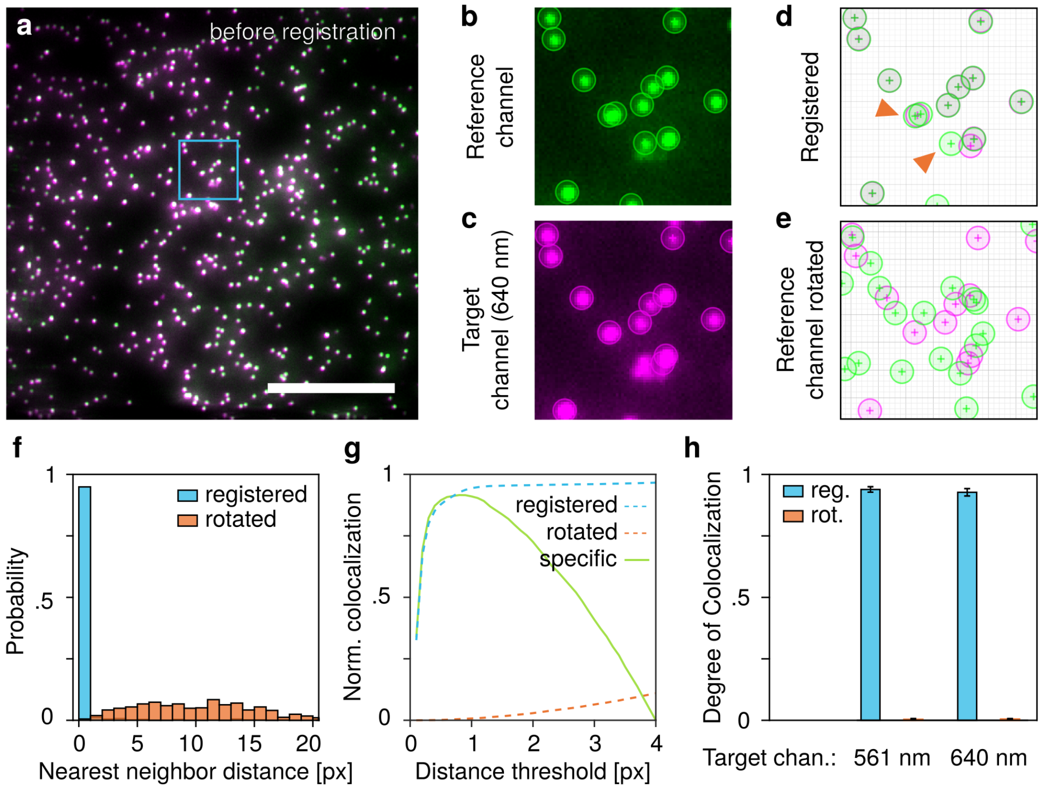

Figure 3.11 Validation of gSEP pipeline with multispectral beads. a, Raw images of 0.1 µm Tetraspeck beads. Pre-aligned images are shown in false colors. Green: reference channel with excitation at 488 nm, magenta: target channel with excitation at 640 nm. Scale bar: 15 µm. **b,c** Single-channel images for reference and target channel of region highlighted in (a) with overlayed detected emitter positions for respective channel (colored crosses and circles). **d,** Emitter positions after channel registration. Arrowheads: Missed emitters in target channel. **e,** Emitter positions after reference channel was rotated by 90°. **f,** Nearest neighbor distance between reference and target channel emitters before and after reference channel rotation for detected emitters from image shown in (a). **g,** Normalized distance-dependent colocalizations before and after reference channel rotation. Specific colocalization was computed according to Equation 3.2. Cutoff distance T for dataset shown: 1 px. **h,** Degree of colocalization for colocalization analysis between reference and green or red target channel before and after reference channel rotation. 32 images with 1.98×10^4 reference emitters total. Data was analyzed jointly with Felix Braun, raw data was recorded by me.

to their nearest neighbor in the target channel of one pixel or less (Fig. 3.11f). However, it was also observed that a small fraction of signals that could be identified by eye had not been successfully localized (Fig. 3.11d, orange arrowheads).

To estimate the amount of non-specific colocalization, we compared the target image to a reference image rotated by 90°, which led to a low degree of overlap between reference and target emitter positions (Fig. 3.11e). Consequently, colocalization analysis resulted in larger nearest neighbor distances of $\sim$10 pixel on average (Fig. 3.11f). Those data were then used to identify the distance cutoff value T, which was subsequently used for computing the degree of colocalization (Fig. 3.11g). The final colocalization analysis with a distance cutoff T of 1.0 pixels between the reference and the 640 nm target channel resulted in a degree of colocalization of $93 \pm 1\,\%$ and a non-specific degree of colocalization after rotation of the target channel of $0.5 \pm 0.2\,\%$. Similar degrees of colocalization were obtained after repeating the same analysis with the 561 nm excitation reference channel ($94 \pm 1\,\%$ specific vs. $0.4 \pm 0.2\,\%$ non-specific).

Based on the shown validation experiments, I conclude that the updated version of `DOLanalysis` produces reliable labeling efficiency estimates across a range of different imaging conditions.

3.6 Global performance of SNAP$_f$-tag and HaloTag

In the previous chapter, I described how an initial version of a workflow for measuring labeling efficiencies on the single-molecule level could be improved by optimizing emitter localization. I then benchmarked the updated workflow with simulations as well as experimental data and validated that the newly established `DOLanalysis` framework provided accurate measurements of labeling efficiencies for protein tag-based labeling. In addition, I established a protocol for semi-automated data acquisition, which is essential to increase throughput and therefore provides a basis to extend the investigation of labeling efficiencies toward additional conditions and parameters.

3.6.1 Influence of ligand concentration and incubation time

To study how achieved labeling efficiencies and non-specific ligand binding in samples depends on ligand concentration and incubation time, I set out to perform a systematic evaluation of labeling conditions in the epithelial adenocarcinoma cell line H838[180]. H838 cell lines expressing the gSEP and LynG constructs were generated by retroviral transduction in collaboration with Florian Salopiata (DKFZ Heidelberg). For retroviral transduction, we used the Phoenix ampho system and the plasmids pBABE-Lyn-HALO-linker-SNAPf (gSEP) and pBABE-Lyn-GFP (LynG) as described above (Section 3.2.1). This was done to achieve stable genomic integration of the gSEP and LynG transgenes, which helps to reduce the variability of expression levels between experiments and to facilitate semi-automated image acquisition by maximizing the number of transfected cells per sample.

We verified that both cell lines stably expressed either gSEP or LynG after viral transduction and selection with puromycin by qualitative immunoblotting. The full-length expression of both constructs was confirmed by protein detection with antibodies binding to EGFP located at the N-terminus or to His$_6$-tag located at the C-terminus of gSEP and LynG. For both H838 cell lines, signals corresponding to the expected size of gSEP ($\sim$80 kDa) and LynG ($\sim$30 kDa) were observed (Fig. A.2). We therefore concluded that both constructs are fully expressed, which is critical for downstream data analysis since we assume that each position at which a reference (EGFP) signal is detected represents a potential binding site for protein tag ligands. To verify that the expression strength of both constructs was suited for single-molecule detection, I fixed samples from both cell types and assessed their respective EGFP signal density as proxy for gSEP and LynG expression by TIRF microscopy. As shown in Fig. 3.12a,b individual, diffraction-limited signals could be observed in the majority of cells for both cell types.

I proceeded by preparing sets of samples, where I incubated living cells from both H838 cell types with ligands for SNAP_f-tag and HaloTag at varying concentrations and incubation times of 15 min, 30, 60 min, 3 or 16 h. After removal of unbound ligands by repeated washing, samples were fixed with Paraformaldehyde (PFA) (detailed protocols are described in Section 6.2.3). For SNAP_f-tag, the ligand BG-SiR at concentrations from 0.5 nM to 1.25 µM was added, while for HaloTag, the ligand HTL-TMR was added at concentrations from 0.1 nM to 250 nM. In addition, a control without ligand addition was included for each incubation time. BG-SiR and HTL-TMR were used to distinguish between ligands bound to SNAP_f-tag or HaloTag based on their respective excitation wavelengths, which are 561 nm for HTL-TMR and 640 nm for BG-SiR. This allowed me to perform simultaneous measurements of labeling efficiencies for both protein tags within the same sample and in a single image acquisition session. The described labeling conditions were chosen to match a previous experiment performed in Huh7.5 cells. A comparison of labeling efficiencies obtained in different cell lines will be shown in Section 3.6.2.

Samples for 80 different conditions (2 cell types, 5 incubation times, 8 ligand concentrations) were imaged using the semi-automated data acquisition routine described in section 3.4. This allowed me to image $\geq$10 cells per condition by single-molecule TIRF microscopy and with identical settings for all samples. The obtained dataset comprised a total of 1 283 cells. Prior to downstream analysis, cells were classified based on defined quality criteria (section 6.3.2) and 37.1 % of imaged cells were consequently excluded from the dataset. The remaining 807 image sets each comprised of one reference channel image containing EGFP signals, one target channel image containing HTL-TMR signals and one target channel image containing BG-SiR images were then processed by emitter localization with `thunderSTORM` and localization filtering with `DOLanalysis` as described above.

As this was the first dataset recorded with the new semi-automated data acquisition pipeline, I assessed the detected emitter densities in the reference (EGFP) channel to evaluate the reproducibility of generated data. Across the entire dataset, the detected EGFP emitter density varied considerably between individual cells ranging from 0.05 to 1.65 μm^{-2} for H838 gSEP samples and from 0.23 to 1.88 μm^{-2} for H838 LynG samples (Fig. 3.12c). The median detected emitter density of $0.60 \pm 0.29\,\mu\text{m}^{-2}$ (median±SD) for H838 gSEP cells was comparable to Huh7.5 gSEP cells used for previous experiments, while the expression level

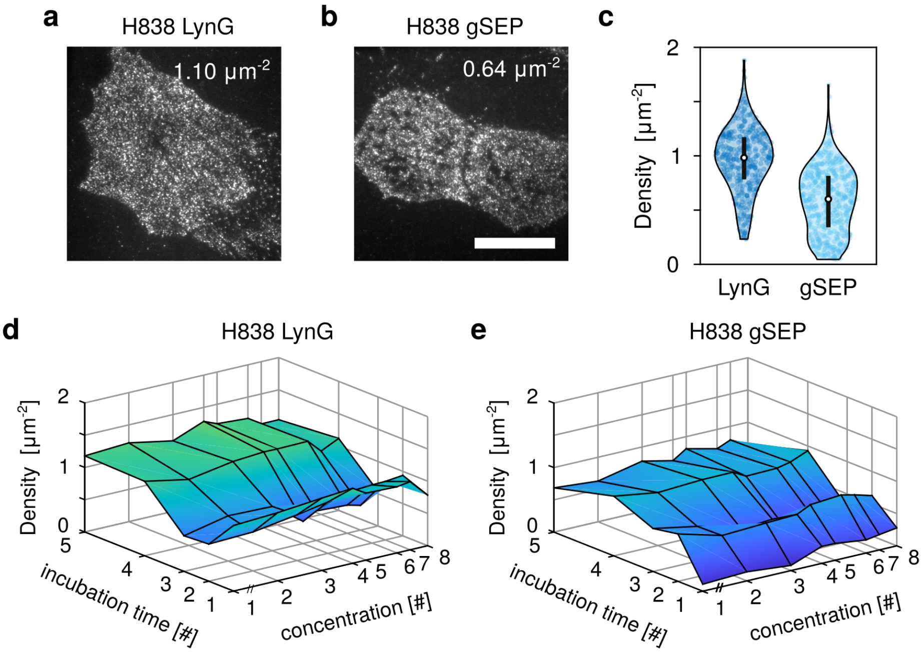

Figure 3.12 Detected emitter density in H838 cell lines in reference (EGFP) channel. **a,b** Representative images of H838-gSEP (a) and H838-LynG (b) cells in the reference (EGFP) channel. Detected EGFP emitter density for individual images are indicated in the top right corner of each image. Scale bar: 20 µm. **c,** Cell-wise detected emitter density across all 40 conditions for H838 gSEP and LynG cell line. **d,e,** Median emitter density for each experimental condition in gSEP (d) and LynG (e) cells. Data from 4-15 (gSEP) and 3-15 (LynG) per condition. 427 (gSEP) and 380 (LynG) cells overall.

for H838 LynG cells ($0.99 \pm 0.29\,\mu m^{-2}$) was ~60 % higher as in the corresponding Huh7.5 LynG cell line (Fig. 3.2c,d).

For both cell types, I also observed variability in detected reference emitter densities across different samples. In H838 LynG, sample-wise detected reference emitter densities varied between $0.30 \pm 0.11\,\mu m^{-2}$ and $1.34 \pm 0.28\,\mu m^{-2}$ as shown in Fig. 3.12d. In H838 gSEP, sample-wise detected reference emitter densities varied between $0.07 \pm 0.04\,\mu m^{-2}$ and $0.91 \pm 0.22\,\mu m^{-2}$ as shown in Fig. 3.12e. In general, samples with identical ligand incubation times tended to exhibit similar detected EGFP emitter densities, but samples with different ligand incubation times varied in their detected emitter densities. For example, systematic deviations towards lower detected emitter densities were observed for samples with H838 gSEP cells at incubation time #1 and #3 (Fig. 3.12d). This observation was unexpected as typical reference emitter numbers should be approximately constant across across different labeling conditions because expression of gSEP and LynG is independent of labeling and only depends on the used cell line.

It is important to note that no systematic dependence of detected reference emitter densities with incubation time or ligand concentrations was observed. Also, even in the samples with the lowest detected emitter density (H838 gSEP, incubation time #1), at least 240 ± 113 reference channel emitters per were still detected per cell (median±SD), which is sufficient for evaluating labeling efficiencies based on single-molecule colocalization analysis. I therefore proceeded by performing the colocalization analysis with `DOLanalysis` as described in Section 3.2.2 to determine the labeling efficiencies for SNAP$_\mathrm{f}$-tag and HaloTag across the full set of ligand concentrations and incubation times.

In the target channels, which contained signals from the protein tag ligands BG-SiR and HTL-TMR, I expected to observe a systematic dependence of emitter numbers on the applied labeling conditions. In the case of H838 gSEP cells, increased binding of ligands gSEP upon increasing incubation times or ligands concentrations should lead to an increased colocalization between emitters detected in the reference and target channels. Consequently, colocalization analysis between the reference channel and the target channels should reveal a dependence of measured labeling efficiencies on ligand concentration, as well as incubation time.

Indeed, I observed a rapid increase in the measured labeling efficiency in H838 gSEP for labeling of HaloTag with HTL-TMR at ligand concentrations as low as 0.1 nM and incubation times as short as 15 min (Fig. 3.13a). For an incubation time of 15 min an average increase in labeling efficiency from 16 ± 6 % to 38 ± 7 % was observed upon increasing HTL-TMR concentrations from 0.1 to 1 nM, but further increasing ligand concentration did not lead to increased labeling efficiencies (Fig. 3.13b). At a ligand concentration of 0.1 nM, a continuous increase of measured labeling efficiencies was observed for incubation times of up to 1 h.

After 1 h incubation at a concentration of 1 nM the achieved labeling efficiency was 30 ± 7 %. Further increasing the incubation time to 16 h of labeling, resulted in further moderate increase in labeling efficiency to 42 ± 10 % (Fig. 3.13c). As expected, no specific labeling was detected if no HTL-TMR had been added to samples (Fig. 3.13b, dashed line), or in cells expressing LynG instead of gSEP (Fig. A.6a). Incubation times, where below-average reference emitter densities were observed previously (15 min and 1 h), also tended to show systematic trends towards lower labeling efficiencies.

Importantly, and in contrast to the general assumption no labeling efficiency beyond 50 ± 12 % was observed across any of the tested conditions. In addition, the observed saturation behavior of the labeling efficiency, suggested the maximally possible labeling

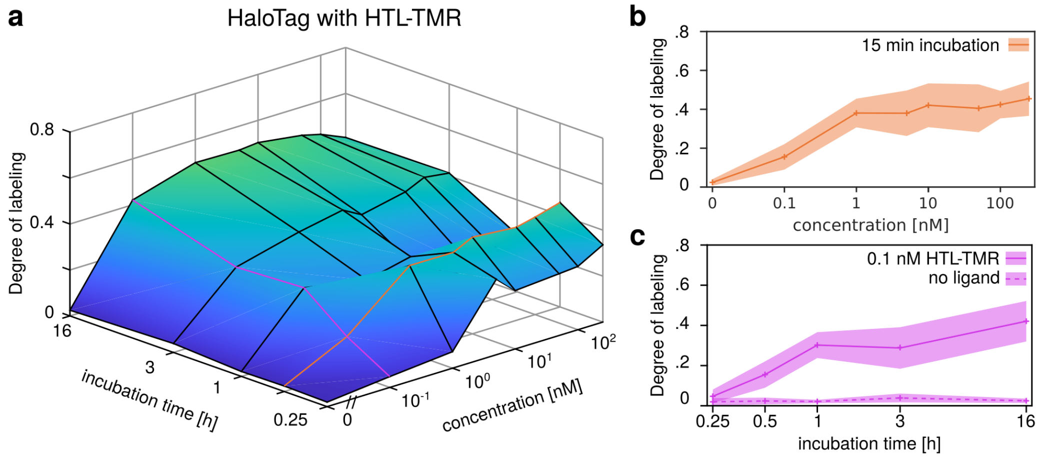

Figure 3.13 Labeling efficiency of HaloTag with HTL-TMR in H838 cells under varying labeling conditions. a, Median DOL for indicated conditions. Colored lines indicate conditions for data shown in b,c. **b,** Median±SD DOL at different ligand concentrations and 15 min incubation time. **c** Median±SD DOL at 0.1 nM ligand concentration and different ligand incubation times. Dashed magenta line: Median±SD DOL without ligand addition. 4-15 cells per condition.

efficiency for labeling of HaloTag with HTL-TMR in living cells had been reached as early as 15 min at a ligand concentration of 1 nM.

For labeling with BG-SiR, a more gradual increase of the measured labeling efficiency along both labeling parameters was observed. While for combinations of ligand concentrations $\leq$1 nM and incubation times $\leq$1 h no specific labeling was observed, combinations of ligand concentrations $\geq$25 nM and incubation times $\geq$3 h led to high labeling efficiencies $>$50 % (Fig. 3.14a), which was consistently higher than the maximum labeling efficiencies achieved for HaloTag labeling with HTL-TMR. At the same time, the total number of labeling conditions under which labeling efficiencies seemed to have been saturated were lower for SNAP$_f$-tag labeling with BG-SiR compared to HaloTag labeling with HTL-TMR. This is in line with the reported reactivity of SNAP$_f$-tag and HaloTag, where for HaloTag efficient labeling at shorter incubation times and lower ligand concentrations was observed[70].

At a constant incubation time of 30 min, observed labeling efficiencies continuously increased up to a ligand concentration of 250 nM, where a labeling efficiency of around 49 ± 14 % was reached. At concentrations beyond that, the labeling efficiency did not seem to increase further, i.e., saturated labeling had been reached (Fig. 3.14b). At a constant BG-SiR concentration of 5 nM, increasing labeling efficiencies were observed with increasing incubation time (Fig. 3.14c). As already observed for HaloTag labeling with HTL-TMR, the cell-to-cell variability observed for labeling of SNAP$_f$-tag with BG-SiR tended to be higher at incubation times ≥3 h. Again, no systematic dependence of labeling efficiencies measured without ligand addition was observed (Fig. 3.14c, dashed line). Finally, labeling efficiencies measured in samples with incubation times of 15 min and 1 min again systematically deviated from measurements at adjacent conditions.

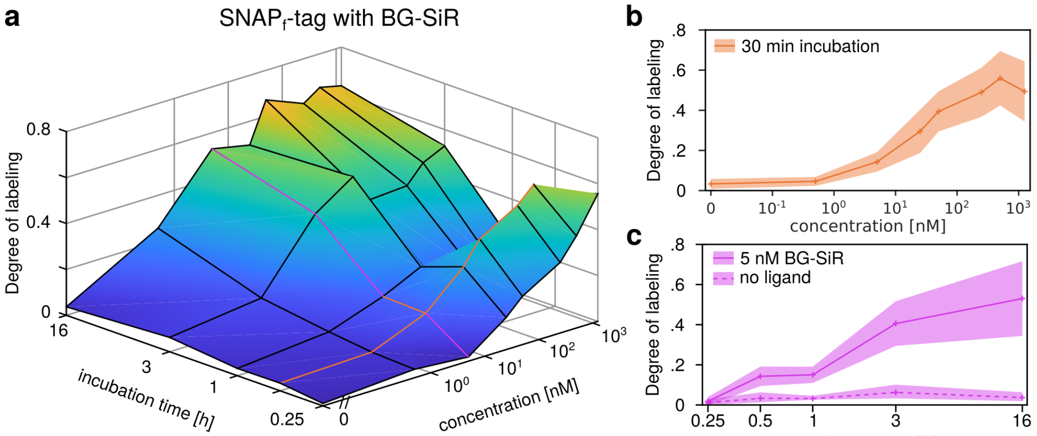

Figure 3.14 Labeling efficiency of SNAP$_f$-tag with BG-SiR in H838 cells under varying labeling conditions. a, Median DOL for indicated conditions. Colored lines indicate conditions for data shown in (b,c). **b** Median±SD DOL at different ligand concentrations and 15 min incubation time. **c,** Median±Standard deviation (SD) DOL at 5 nM ligand concentration and different ligand incubation time. Dashed magenta line: Median±SD DOL without ligand addition. 4-15 cells per condition

After having determined the specific labeling efficiency for SNAP$_f$-tag and HaloTag in live H838 cells using the gSEP construct, I assessed the tendency for non-specific binding for the two ligands HTL-TMR and BG-SiR in H838 LynG cells. In these cells, which lack specific binding sites for HTL-TMR and BG-SiR, substantially lower emitter densities as compared to H838 gSEP were detected in both target channels (Fig. 3.15). However, a weak dependence of emitter densities on incubation time and ligand concentration was observed. For example, at a concentration of 1.25 µM, the non-specific attachment of BG-SiR resulted in detected emitter densities increasing from ∼0.22 to ∼0.74 μm^{-2} at incubation times of 15 min and 16 h, respectively. For HTL-TMR, non-specific attachment in H838 cells was only observed at concentrations >50 nM and at an incubation time of 16 h. At ligand concentrations below 50 nM, no significant dependence of non-specific ligand deposition on incubation time was observed for either ligand.

In principle, colocalization analysis in LynG cells allows to determine an apparent labeling efficiency caused by random colocalization of non-specific binding of ligands to cells. However, to be able to compare the apparent labeling efficiency in LynG cells to the labeling efficiency measured in gSEP cells, both cell types should exhibit approximately

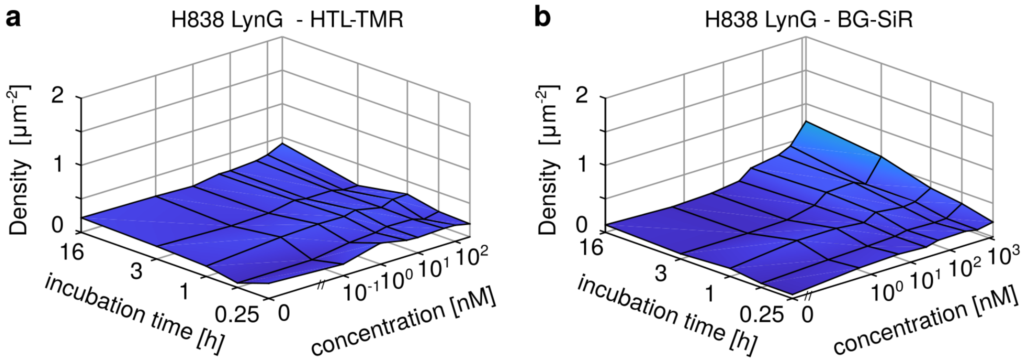

Figure 3.15 Influence of labeling conditions on non-specific ligand deposition in H838 LynG cells. Median detected target channel emitter densities after localization filtering in H838 or LynG cells under variable labeling conditions with HTL-TMR (a) or BG-SiR (b). Data from 3-15 cells per condition.

equal expression levels. Since H838 LynG cells exhibited, ~60 % stronger expression, the obtained labeling efficiencies should not be directly compared. Fig. A.6a,b shows the obtained labeling efficiencies across all tested combinations of ligand concentration and incubation time to provide to an upper estimate of the fraction of the labeling efficiencies measured in H838 gSEP cells due to non-specific deposition of tag ligands. For HTL-TMR a maximum apparent labeling efficiency of 8 ± 1 % was measured at the highest tested ligand concentration of 250 nM and an incubation time of 16 h. In the case of labeling with BG-SiR, the highest apparent labeling efficiency (27 ± 3 %) was also observed at the maximum incubation time of 16 h and at the maximum tested ligand concentration of 1.25 µM.

3.6.2 Labeling efficiency across different cell lines

In the previous section, I showed that the labeling efficiency for $\mathrm{SNAP_f}$-tag and HaloTag in living H838 cells depends on the specific labeling conditions. I also showed that HaloTag exhibited saturated labeling at lower ligand concentrations and incubation times than $\mathrm{SNAP_f}$-tag. Finally, I showed that in H838 cells, the saturation labeling efficiency for $\mathrm{SNAP_f}$-tag with BG-SiR was higher than the saturation labeling efficiency for HaloTag with HTL-TMR.

At this point it was unclear whether the observed performance for both tags was due to specific properties of the tags, the ligands used or the cell line in which measurements were performed. For this reason, I decided to address these questions by comparing the results obtained in H838 cells described above with measurements in Huh7.5 and HeLa cell lines. For establishing HeLa gSEP and HeLa LynG, I again collaborated with Florian Salopiata, who performed stable transductions of wild-type HeLa cells as described above for H838 cells (Section 3.6.1). In contrast to H838 cells, HeLa cells required selection of a sub-population of gSEP and LynG cell clones with suited expression levels. I consequently performed Fluorescence-activated cell sorting (FACS) on both cell lines after stable transduction (see Section 6.1.3). After sorting, HeLa gSEP exhibited a median detected emitter density in the reference (EGFP) channel of $0.40 \pm 0.11\,\mu\mathrm{m}^{-2}$ (median±SD), while HeLa LynG cells exhibited a median detected reference emitter density of $0.21 \pm 0.19\,\mu\mathrm{m}^{-2}$. Full-length expression of gSEP and LynG in HeLa cells after FACS was again confirmed by immunoblotting (Fig. A.2). Labeling efficiency measurements under identical conditions as described in Section 3.6.1 were then performed by Wioleta Chmielewicz. A third dataset recorded in Huh7.5 cells[181] under identical labeling conditions was obtained from previous work by Siegfried Haenselmann[174].

To ensure that labeling efficiencies could be compared across data obtained in different cell lines, I first manually curated both datasets from HeLa and Huh7.5 cell lines comprised of 1605 (HeLa) and 876 (Huh7.5) individual image sets according to the same principles used for curation of datasets recorded in H838 and described in section 6.3.2. In the HeLa cell dataset, 54.6 % of image sets passed quality control, while in the Huh7.5 dataset, which was acquired manually and not with the semi-automated acquisition pipeline, 74.8 % of image sets passed quality control. I proceeded by processing all three datasets by first performing emitter localization with `thunderSTORM` and all further analysis steps with `DOLanalysis` using identical settings as in the H838 dataset describe above.

The additional datasets for SNAP$_f$-tag labeling with BG-SiR and HaloTag labeling with HTL-TMR in Huh7.5 and HeLa cells exhibited qualitatively similar features as the dataset acquired in H838 and described in detail in the previous section. To compare datasets recorded in different cell lines on a global scale, I fitted the observed experimental data with a model describing the underlying labeling reaction. For this, I assumed a simplified labeling reaction where unbound ligand (L) reacts with protein tags (P) to form a stable complex (PL) in an irreversible reaction:

$$\mathrm{P} + \mathrm{L} \longrightarrow \mathrm{PL}$$

In addition, I assumed that the concentration of ligand over time can be considered to be constant during the labeling reaction. This assumption is in line with properties of the experimental system which has been characterized by comparing molecule numbers for free ligand and gSEP based on both, the reported volume of a HeLa cells[182] and the segmented area of cells (this work) assuming a density of gSEP molecules of 1 µm^{-2} (see Fig. 3.12c and Fig. A.4b). In both cases, even at the lowest ligand concentration (0.1 nM) used in this study, the total amount of ligand molecules in a sample is 1×10^3-1×10^5-fold higher than the overall number of gSEP molecules on the surface of cells across all cell in a sample. These two assumptions allowed me to consider the reaction to be of pseudo first-order:

$$LE(t) = \alpha(1 - e^{-[L]k_{app}t}) \tag{3.5}$$

Here, α is a scaling factor to account for saturated labeling at $\leq$100 % labeling efficiency, k_{app} is the apparent rate constant, $[L]$ is the ligand concentration and t is the incubation time. Fitting this model to labeling efficiencies obtained at different $[L]$ and t in different cell lines enabled me to compare labeling of SNAP$_f$-tag or HaloTag with respect to their estimated saturation labeling efficiency (α) and labeling kinetics (k_{app}). As shown in Fig. 3.16, the model was able to describe experimental data recorded in H838 cells for labeling of SNAP$_f$-tag with BG-SiR (Fig. 3.16a,c) and HaloTag with HTL-TMR (Fig. 3.16b,d) with respect to both, the kinetics of the labeling reaction, as well as the labeling efficiency at which saturated labeling occurred.

The general trends for labeling efficiencies with SNAP$_f$-tag and HaloTag in HeLa and Huh7.5 cells mirrored the features described above for H838 cells (Fig. A.5). Labeling of HaloTag with HTL-TMR was more efficient as saturated labeling was achieved at lower ligand concentrations and incubation times. At the same time, the saturation labeling

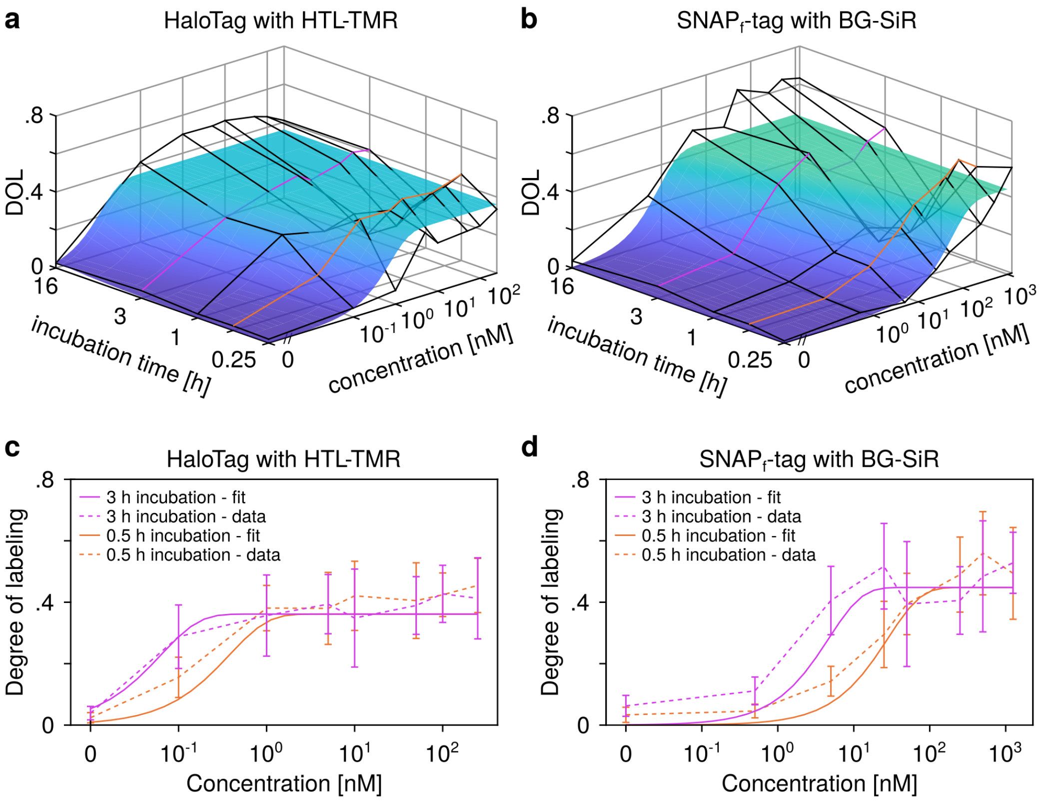

Figure 3.16 Comparison of measured DOLs and predicted DOLs from fit with pseudo-first order model (Equation 3.5). a,b Median measured DOLs in H838 gSEP cells (mesh, same data as in Fig. 3.14a and Fig. 3.13a) and predicted DOL (colored surface) obtained from fit for HaloTag labeling with HTL-TMR (a) and SNAP$_f$-tag labeling with BG-SiR (b). **c,d** Comparison of median±SD measured DOLs (lines) and modeled DOLs (dashed lines) for HaloTag (c) and SNAP$_f$-tag (d) for conditions indicated as colored lines in a,b.

efficiency achieved with SNAP$_f$-tag labeling with BG-SiR appeared to be slightly higher. Incubation times at which measured labeling efficiencies seemed to systematically deviate towards lower values were also observed in Huh7.5 datasets which were acquired without using the newly established semi-automated acquisition pipeline. In contrast to samples in H838 and Huh7.5 cells, labeling efficiencies measured in HeLa cells for both tags tended to exhibit higher variation between cells for an individual condition and between median labeling efficiencies achieved for adjacent conditions (Fig. A.5a,b).

While qualitatively, the performance of SNAP$_f$-tag and HaloTag appeared similar across cell lines, the reaction model fits revealed quantitative differences between cell lines. Overall, for SNAP$_f$-tag labeling with BG-SiR, apparent rate constants between 1.5-8.1 $\times 10^4$ M^{-1} s^{-1} were observed (Table 3.1). Here, k_{app} was similar for labeling in H838 and HeLa cells, while k_{app} in Huh7.5 was 4 times higher. In agreement with the maximum labeling efficiencies observed in experiments for the three cell lines, model fits resulted in the highest saturation labeling efficiency, α of 0.45 ± 0.06 for labeling in H838 cells, while in HeLa and Huh7.5 lower α of 0.36 ± 0.04 and 0.36 ± 0.03 were observed

For HaloTag labeling with HTL-TMR, ~40-70x higher rate constants between 6.3×10^5-$3.7 \times 10^6\ \mathrm{M^{-1}\,s^{-1}}$ were obtained from fitting (Table 3.1). Here, labeling in H838 and Huh7.5 cells led to similar k_{app}, while labeling in HeLa cells led to ~2-4-fold lower k_{app}. Across all three cell lines, α values for HaloTag labeling with HTL-TMR were 2-9 % lower compared to labeling of SNAP$_f$-tag with BG-SiR. Again, the highest α of 0.36 ± 0.04 was observed for labeling in H838, while the lowest α (0.24 ± 0.02) was observed in HeLa cells.

Table 3.1 Labeling efficiencies for HaloTag and SNAP$_f$-tag across cell lines. HaloTag was labeled with HTL-TMR, SNAP$_f$-tag was labeled with BG-SiR. k_{app} and α: Apparent rate constant and scaling factor obtained from fitting experimental DOL values at varying labeling conditions with Equation 3.5. Max. DOL$_{exp}$: Maximum DOL observed in experiments under maximum DOL condition. Raw data in Huh7.5 cells recorded by Siegfried Haenselmann, raw data recorded in HeLa cells recorded by Wioleta Chmielewicz. All data was processed and analyzed by me.

Cell type	k_{app} [$\mathrm{M^{-1}\,s^{-1}}$]	$\alpha * 100$	DOL$_{max,exp}$ [%]	max. DOL cond.
SNAP$_f$-tag				
H838	$(2.0 \pm 1.3) \times 10^4$	45 ± 6	67 ± 14	50 nM,16 h
HeLa	$(1.5 \pm 1.2) \times 10^4$	26 ± 4	41 ± 5	250 nM, 0.5 h
Huh7.5	$(8.1 \pm 4.5) \times 10^4$	36 ± 3	57 ± 14	250 nM, 3 h
HaloTag				
H838	$(1.5 \pm 1.2) \times 10^6$	36 ± 4	50 ± 12	10 nM,3 h
HeLa	$(6.3 \pm 8.5) \times 10^5$	24 ± 4	42 ± 10	5 nM, 3 h
Huh7.5	$(3.7 \pm 2.3) \times 10^6$	29 ± 2	42 ± 6	0.1 nM, 16 h

3.6.3 Alternative fluorescent ligands for labeling in living cells

So far, labeling of SNAP$_f$-tag and HaloTag was done with the same ligands, BG-SiR and HTL-TMR. By comparing labeling efficiencies across different incubation times and ligand concentrations in three different cell lines, I observed that the cell line in which both tags were expressed influenced labeling kinetics as well as the reachable saturation labeling efficiency. Across all tested cell lines, I observed saturation of ligand binding to both tags at labeling efficiencies ≪100 %, which showed that this behavior was well reproducible across different experiments and influenced only to a small degree by the intracellular environment in H838 cells. At this point, limited reactivity of SNAP_f-tag and HaloTag towards BG-SiR and HTL-TMR represented a possible explanation for limited labeling efficiencies.

To determine if improved labeling could be obtained with alternative ligands, I tested if exchanging the fluorophores conjugated to BG or HTL led to differences in labeling efficiency. This was done by first swapping TMR and SiR as fluorophores conjugated to BG and HTL to produce BG-TMR and HTL-SiR and second by extending the set of tested fluorophores with the TMR and SiR derivatives, Janelia Fluor (JF)549 and JF646. These fluorophores were recently developed by Grimm et al. and were reported to exhibit improved photophysical properties while maintaining cell permeability[183].

For all three additional ligands, the labeling efficiencies at two different concentrations were measured for SNAP_f-tag and HaloTag. The tested concentrations varied between tags to reflect the difference in reactivity previously observed for SNAP_f-tag and HaloTag. The first concentration was chosen such that it was around the minimum saturation concentration for SNAP_f-tag with BG-SiR or for HaloTag with HTL-TMR. The second concentration was chosen to in a way that saturated labeling with BG-SiR and HTL-TMR was achieved. As Huh7.5 cells exhibited a lower expression strength for gSEP compared to H838 cells, the experiments were performed in Huh7.5 gSEP and Huh7.5 LynG cells to minimize the influence of emitter detection and subsequent density correction.

For HaloTag, all three additional ligands exhibited a dependence of the achieved labeling efficiency on ligand concentration (Fig. 3.17a). The smallest difference in achieved labeling efficiency between labeling at 1 nM and 10 nM ligand concentration was observed for HTL-SiR, indicating that the affinity of HTL-SiR for HaloTag was highest among the tested ligands. The labeling efficiency achieved at labeling with HTL-SiR at 100 nM ligand concentration was comparable to the saturation labeling efficiency previously observed for HTL-TMR in Huh7.5 cells (Fig. A.5).

For the two JF-ligands, larger differences in labeling efficiency between labeling at 1 and 10 nM ligand concentration were observed. Here, labeling with 1 nM ligand resulted in very low labeling efficiencies of 2 ± 1 and 5 ± 3 % for HTLJF549 and HTL-JF646, respectively. This indicates that the affinity of HTL-JF ligands for binding to HaloTag was reduced compared to HTL-TMR and -SiR. The labeling efficiencies achieved at 100 nM ligand concentration for HTLJF549 and HTL-JF646 of 26 ± 10 % and 39 ± 16 % were again similar to the saturation labeling efficiency achieved with HTL-TMR. The non-specific emitter densities detected after labeling with HTL-SiR and HTL-JF646 did not exceed background

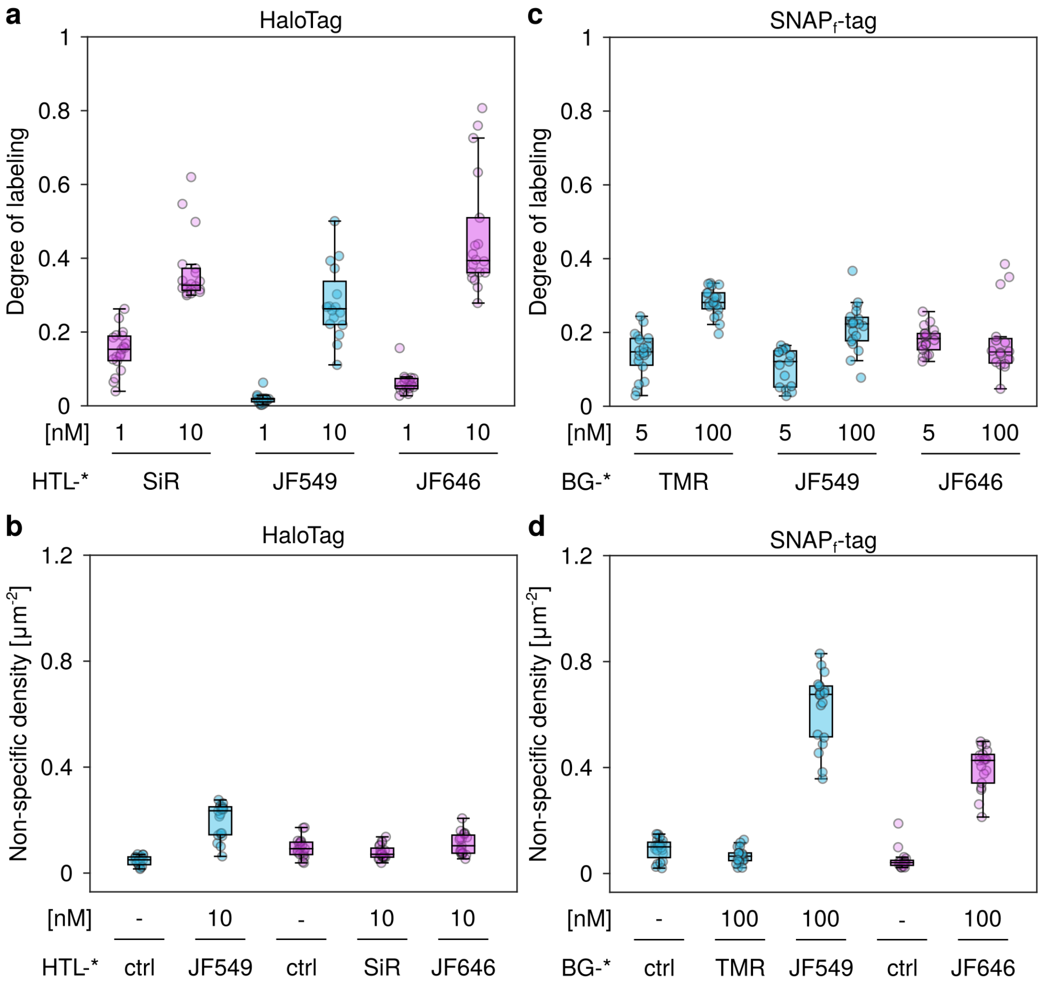

Figure 3.17 Labeling efficiency and non-specific background for SNAP$_f$-tag and HaloTag with cell permeable ligands. Live Huh7.5 gSEP or LynG cells were labeled for 30 min with indicated ligands at indicated concentrations to determine labeling efficiencies (a,b) and non-specific labeling density (c,d). Circles represent individual cells, 15-20 cells per condition. Individual plots are color-coded according to excitation wavelength. Cyan: 561 nm, magenta: 640 nm. Measurements were performed by Felix Braun. All data was processed and analyzed by me.

levels indicating highly specific labeling for these two ligands (Fig. 3.17b). For HTL-JF549, an increased background emitter density of $0.24 \pm 0.06\,\mu\mathrm{m}^{-2}$ was observed.

For SNAP$_\mathrm{f}$-tag, all three additional ligands showed reduced labeling efficiencies compared to BG-SiR at both tested concentrations (Fig. 3.17c). While for BG-TMR, differences of $\sim$7 % were observed, differences of $>$10 % at both tested ligand concentrations were observed for BGJF549. For BGJF646, almost no difference in measured labeling efficiency between both tested ligand concentrations was observed and the labeling efficiency at 100 nM ligand concentration was $\sim$50 % of the labeling efficiency achieved with BG-SiR at the same ligand concentration and incubation time. The non-specific detected emitter density in Huh7.5 LynG cells for BG-TMR was comparable to background levels. At the same time, the detected non-specific emitter density was strongly increased for both BG-JF ligands (Fig. 3.17d). For BG-JF549, the background emitter density in Huh7.5 LynG cells was increased $>$6-fold over the background emitter density in control cells. For BG-JF646, a slightly lower non-specific emitter density in LynG cells, $\sim$3-fold over the background density in corresponding control cells was observed. This indicated that both, BG-JF549 and BG-JF646 at 100 nM ligand concentration a strong non-specific labeling under the tested conditions

To summarize, I observed that different ligands for SNAP$_\mathrm{f}$-tag and HaloTag exhibited different labeling behavior. HTL-JF549 and HTL-JF646 were identified as promising alternative ligands for HaloTag, while BG-TMR was shown to perform slightly worse than BG-SiR in earlier experiments. Overall, the comparison of labeling efficiencies achieved with three alternative fluorescent ligands for each tag suggests that while labeling efficiencies seem to depend on the applied ligand to some degree, none of the ligands achieved stoichiometric labeling of either tag.

3.6.4 SNAP$_\mathrm{f}$-tag and HaloTag activity after chemical fixation

In the previous sections, I examined how the cell line in which SNAP$_\mathrm{f}$-tag and HaloTag reacted with their respective ligands influenced the achieved labeling efficiencies and which influence different ligands had on measured labeling efficiencies. Both of these comparisons were performed with labeling in living cells, i.e. were limited to cell permeable ligands and live-cell compatible labeling conditions. However, several advanced imaging techniques, such as correlative light and electron microscopy or super resolutions fluorescence microscopy approaches such as direct optical reconstruction microscopy (dSTORM) or DNA-PAINT are incompatible with living cells or require special, cell impermeable labels. In addition, fixation prior to labeling of protein tags could potentially be beneficial for structural preservation of samples since the amount of buffer exchanges on living cells is minimized. Another possible advantage of labeling post fixation is that non-specific deposition of protein tag ligands could be reduced in the changed chemical environment of fixed cells. For this reason, I addressed the question whether SNAP$_f$-tag and HaloTag retain their ability to bind their respective ligands after chemical fixation and whether labeling post-fixation led to changes in observed non-specific labeling.

Initial experiments with Glutaraldehyde (GA) as primary fixative showed that the amount of autofluorescence generated during fixation was prohibitive to single-molecule imaging with the required SNR. For this reason, no further attempts to implement a sample preparation protocol with GA as fixative were undertaken. In contrast to fixation with GA, fixation with PFA resulted in no noticeable increase in autofluorescence and the previously applied data analysis approach with `thunderSTORM` and `DOLanalysis` could be applied without the need to adjust processing parameters. To determine whether chemical fixation with PFA impacted the functionality of SNAP$_f$-tag or HaloTag, I performed a direct comparison of labeling efficiencies achieved across different ligand concentrations for labeling with BG-SiR and HTL-TMR pre and post fixation. Possible differences in non-specific ligand deposition pre or post fixation were addressed in a second set of measurements in HeLa LynG cells.

Before assessing labeling efficiencies pre- and post-fixation, I briefly validated that detected EGFP reference emitter were comparable for samples labeled before and after fixation. For HeLa gSEP, reference emitter densities of $0.30 \pm 0.08\,\mu\text{m}^{-2}$ for the pre-fixation labeling and $0.31 \pm 0.08\,\mu\text{m}^{-2}$ for the post-fixation labeling dataset were observed. For HeLa LynG cells, a reference emitter density of $0.40 \pm 0.10\,\mu\text{m}^{-2}$ was detected for the pre- and a density of $0.28 \pm 0.08\,\mu\text{m}^{-2}$ was detected for the post-fixation labeling dataset.

In line with previous experiments in HeLa cells (Section 3.6.2), specific labeling after 30 min labeling in living cells was observed for SNAP$_f$-tag with BG-SiR at ligand concentrations $\geq$5 nM. Labeling efficiencies with increasing ligand concentrations up to a concentration of 250 nM where a saturation labeling efficiency of $\sim$29 % had been reached (Fig. 3.18a). For labeling of SNAP$_f$-tag post-fixation, a comparable concentration-dependent behavior was observed. However, the saturation labeling efficiency for post-fixation labeling with BG-SiR was 2.3-fold higher as for pre-fixation labeling ($\sim$69 %).

For HaloTag labeling with HTL-TMR, labeling prior to fixation resulted in saturated labeling after 30 min for a ligand concentration of at least 1 nM (Fig. 3.18b). At concen-

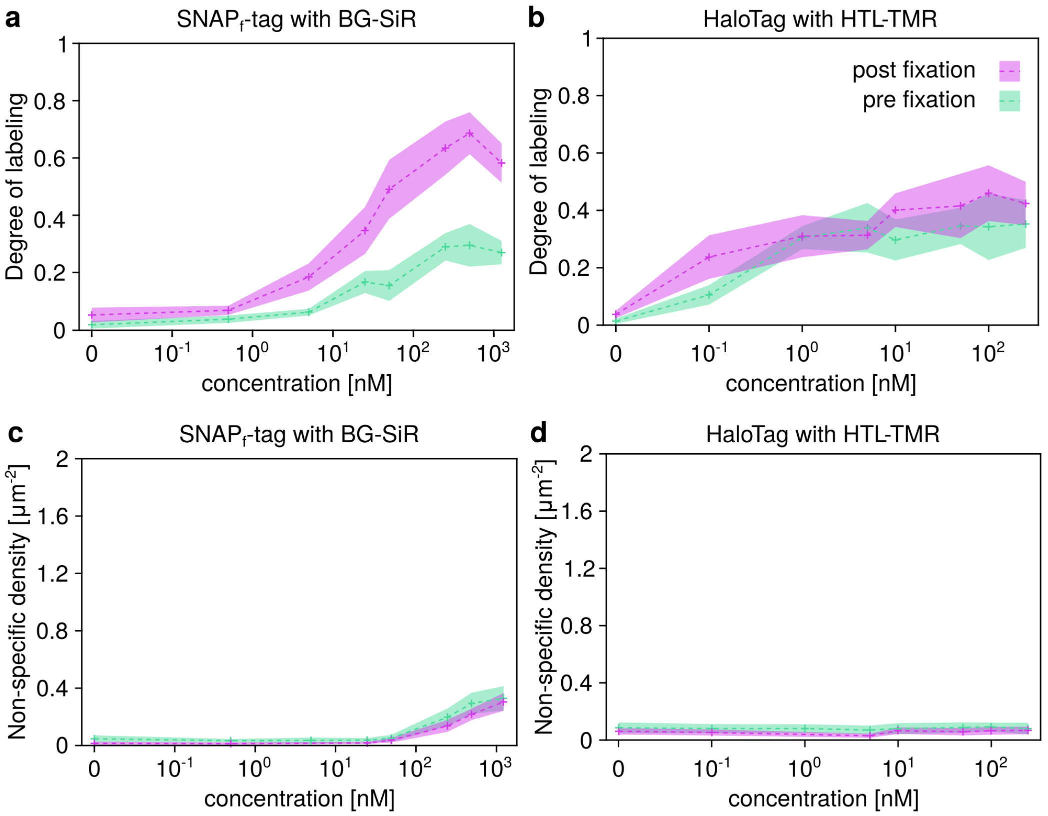

Figure 3.18 Labeling efficiency for labeling post fixation. HeLa gSEP and LynG cells were labeled with varying concentrations of BG-SiR (a,c) and HTL-TMR (b,d) for 30 min pre- (magenta) or post-fixation (green) with 3.7 % PFA. **a,b** DOL measured in gSEP cells at indicated ligand concentrations. **c,d** Unspecific labeling measured as detected point density in LynG cells at indicated ligand concentrations. median±SD (crosses, shaded bands) from 6-20 cells per condition.

trations ≥10 nM, labeling efficiencies of 31-35 % were observed. This behavior was again highly comparable to previous experiments in HeLa cells. In contrast to the observations made for SNAP$_f$-tag, labeling of HaloTag with HTL-TMR post-fixation resulted in smaller differences to pre-fixation labeling. At a ligand concentration of 0.1 nM, a labeling efficiency of 11 ± 3 % for labeling pre-fixation and 24 ± 6 % for labeling post-fixation was observed. The achieved saturation labeling efficiency at ligand concentrations ≥10 nM was increased ∼1.3-fold (40-46 %).

The non-specific ligand deposition for both, BG-SiR and HTL-TMR remained virtually unchanged when switching from pre- to post-fixation labeling. As observed before, BG-SiR exhibited a tendency for non-specific binding in HeLa cells at concentrations ≥100 nM.

3.6.5 AF647 ligands for labeling in fixed cells

As described in section 2.3, quantitative SMLM is a popular approach for emitter counting microscopy and SMLM in fixed samples has been realized with a large variety of different organic fluorophores[184]. However, Alexa Fluor (AF)647 is arguably the most widely used fluorophores in SMLM due to its high brightness and propensity to reside in a metastable dark state[185]. A recent study determined the effective labeling efficiency for labeling of SNAP$_f$-tag and HaloTag with AF 647-conjugated ligands using the nuclear pore complex as calibration target[92].

In order to compare labeling efficiencies obtained by this SMLM-based approach with the gSEP approach established here, I performed labeling efficiency measurements with BG-AF647 and HTL-AF647 under conditions identical to those used by Thevathasan et al.. In contrast to fluorophores used for all experiments described so far, AF647 is not cell permeable, i.e. it can not pass the plasma membrane of living eukaryotic cells. AF647 conjugates therefore require chemical fixation and permeabilization of the plasma membrane prior to labeling. For this reason, I used a sample preparation protocol, which included pre-fixation and permeabilization with Triton X-100 (TX-100) prior to PFA fixation and labeling. To maintain comparability with measurements by Thevathasan et al., I also included a quenching step with NH_4Cl, a blocking step after fixation and added Dithiothreitol (DTT) as well as Bovine serum albumin (BSA) to the labeling buffer (see section 6.2.3). Labeling with ligands was performed for 2 h at room temperature.

Samples prepared according to this protocol were then imaged by TIRF microscopy and data was processed with `thunderSTORM` and `DOLanalysis` as described for previous experiments. Both, BG-AF647 and HTL-AF647 were able to bind their respective tags and achieved labeling efficiencies between 22 and 69 % depending on protein tag and ligand concentration (Fig. 3.19a). For a control measurement in unlabeled wt cells, I observed a median labeling efficiency of 0 %, i.e. no labeling, showing that the altered sample preparation protocol did not lead to increased autofluorescent background levels which interfered with the gSEP measurements. For BG-AF647, a labeling efficiency of 69 ± 18 % (median±SD) was achieved after labeling at 1 µM ligand concentration. At 10-fold lower ligand concentration (100 nM), the labeling efficiency remained unaltered (69 ± 11 % indicating the saturation of active SNAP$_f$-tag had occurred already at this lower concentration. This is in line with my previous observations for SNAP$_f$-tag labeling with BG-SiR (Fig. 3.16 and Fig. A.5), where saturation was estimated to have occurred under the same labeling conditions. Interestingly, the labeling efficiency achieved with BG-AF647 here was identical to the labeling efficiency achieved with BG-SiR when labeling was performed after fixation with PFA (Fig. 3.18).

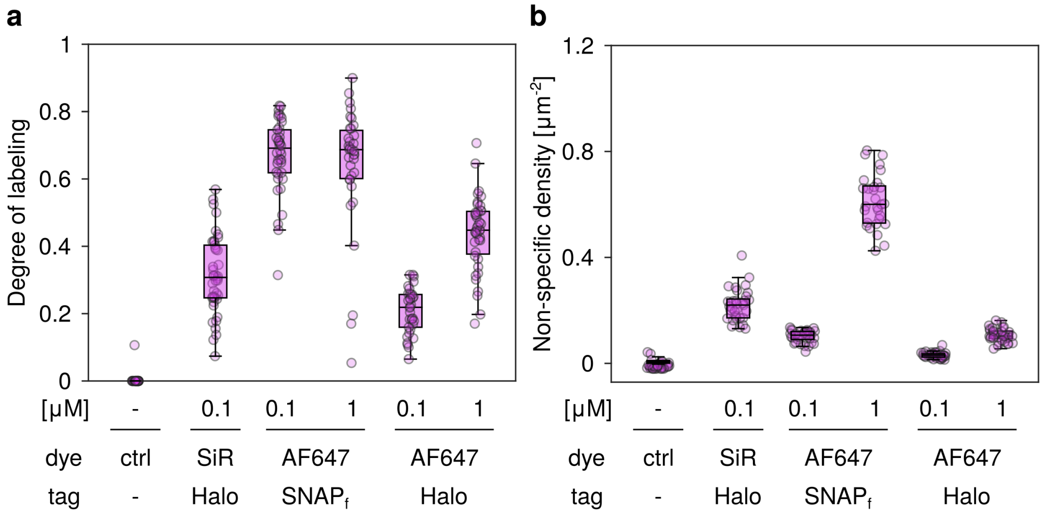

Figure 3.19 Labeling efficiency with AF647 ligands. Huh7.5 gSEP and LynG cells were fixed as described in section 6.2.3 and labeled with AF647 conjugates at indicated concentrations for 120 min. Labeling efficiency (a) and unspecific background (b) for SNAP$_f$-tag and HaloTag AF647 ligands were determined using gSEP (a,b) or LynG cell (c,d). Control: unlabeled Huh7.5 wt cells. Circles represent individual cells, 29-42 cells per condition.

For HTL-AF647, labeling efficiencies of 22 ± 7 % at 100 nM and 45 ± 11 % at 1 µM were observed, indicating that saturation of binding had not been reached after labeling for 2 h at 100 nM HTL-AF647. This is in contrast to labeling with other HaloTag-ligands in fixed and living cells, where saturated labeling was observed at ligand concentrations as low as ≤1 nM if labeling was performed for ≥1 h (Fig. 3.16 and Table 3.1). In contrast to SNAP$_f$-tag labeling with BG-AF647, the labeling efficiency at 1 µM HTL-AF647 observed for HaloTag was higher than the saturation level observed in fixed cells labeled with HTL-TMR (Fig. 3.18a). As a dependence of achieved labeling efficiencies on the specific ligand used for labeling was previously also observed for live-cell compatible HaloTag ligands (Section 3.6.3), the diverging behaviors for HTL-TMR and HTL-AF647 in fixed cells support the hypothesis that for HaloTag, the achievable labeling dependency depends on the molecular structure of the ligand.

Despite the altered labeling protocol, which included a blocking step and addition of BSA to the labeling buffer, it could not be excluded that non-specific deposition of ligands contributed to the measured labeling efficiencies. I therefore evaluated the tendency of BG-AF647 and HTL-AF647 for non-specific label deposition in LynG cells not expressing SNAP$_f$-tag or HaloTag. As for the labeling efficiency, differences between both AF647 ligands were observed with respect to their tendency for non-specific label deposition (Fig. 3.19b). Both ligands showed non-specific emitter densities above a corresponding negative control without added ligand. A concentration-dependent (sublinear) increase in non-specific emitter density was also observed for both types of ligand. Labeling with 1 µM BG-AF647 resulted in non-specific emitter densities of $0.6 \pm 0.1\,\mu\text{m}^{-2}$ (median±SD), while labeling with HTL-AF647 at 1 µM, resulted in a non-specific emitter density of $0.10 \pm 0.03\,\mu\text{m}^{-2}$. This translated to apparent labeling efficiencies of 9 % for BG-AF647 and 1 % for HTL-AF647 at 1 µM concentration. At 100 nM ligand concentration, non-specific detected emitter densities of 0.11 ± 0.02 and $0.03 \pm 0.01\,\mu\text{m}^{-2}$ were observed for BG-AF647 and HTL-AF647 respectively. This observation once again supported the hypothesis that BG ligands for SNAP_f-tag exhibit a stronger tendency for non-specific binding in cells. In contrast to previous observations with different live-cell compatible HaloTag ligands, labeling with HTL-SiR under conditions optimized for labeling with AF647 ligands resulted in a slightly elevated non-specific emitter density of $0.22 \pm 0.06\,\mu\text{m}^{-2}$ leading to an apparent labeling efficiency of $\sim$3 % in LynG cells.

To summarize, I observed that for HTL-AF647, a higher ligand concentration led to improved labeling efficiency while still showing low non-specific labeling meaning that labeling at 1 µM would be the preferred approach. For BG-AF647, a ligand concentration of 1 µM resulted in elevated non-specific label deposition while no benefit for the achieved labeling efficiency could be observed. If labeling is to be performed with BG-AF647, an intermediate ligand concentration would be the preferred approach. Overall, both ligands show good (BG-AF647) or excellent (T̋LAF) specificity and labeling efficiencies en-par or higher than the previously tested live-cell compatible ligands which mean that both are well suited for emitter counting microscopy experiments if labeling is performed post fixation.

3.7 Discussion

The work presented in the course of this chapter established the staining efficiency probe gSEP and the data analysis framework `DOLanalysis` as a tool to measure the labeling efficiency for self-labeling protein tags based on single-molecule colocalization analysis. The gSEP approach was then used to for an in-depth analysis of the labeling behavior of SNAP_f-tag and HaloTag. In the following, I will provide a technical discussion of the results described above. General implications of the observations made in this chapter in the context of protein counting by fluorescence microscopy will be discussed in Part III.

In section 3.4, I used simulations to show that emitter localization with `thunderSTORM`, a software package typically used for SMLM, improved the emitter localization efficiency across all tested conditions. I subsequently showed that post detection filtering of detected emitters based on the sub-pixel localization parameter sigma led to marked reductions in the false-positive detection rate and an concomitant increase in the overall detection performance. Filtering of localizations further showed that prior to filtering, localizations with high amplitude occurred in the raw data. As these localizations originate from emitter aggregates, it was crucial to remove these localizations prior to downstream analysis to avoid biased labeling efficiency estimates. As emitter localizations are the basis for all subsequent analysis steps, improved emitter detection can be expected to increase the robustness of `DOLanalysis` in subsequent studies.

The established analysis routine for comparing emitter localization algorithms on the basis of simulated cell images represents a platform, which can readily be used to evaluate additional emitter localization algorithms in the future. In a recent study, Sage et al. provide a comprehensive comparison of different localization algorithms for SMLM, which can serve as a starting point to identify candidate algorithms[175]. For example, the algorithm FALCON was developed to explicitly address the challenges posed by emitter localization at high emitter densities and could therefore lead to further improvements in localization efficiency[186]. It was recently reported that emitter localization using convolutional neural networks led to further improvements in emitter localization efficiency outperforming conventional algorithms across a wide range of detection settings and emitter densities[187,188]. Such approaches therefore represent another promising direction for further optimizing the localization efficiency.

Since gSEP and the associated data analysis principles had only been used in a proof-of-principle study prior to this work, I validated key steps in the data analysis pipeline prior to using gSEP and `DOLanalysis` to assess labeling efficiencies for SNAP$_\mathrm{f}$-tag and HaloTag. In section 3.5, I first demonstrated that with the chosen emitter localization approach based on `thunderSTORM`, detection of emitters in experimental data is highly specific by showing that $>$90 % of emitters detected in the reference channel originate from EGFP molecules contained in the gSEP or LynG constructs. Second, I made use of simulations to show that the necessary correction to account for missed emitters during localization produces valid labeling efficiency estimates at different overall emitter densities and background levels. Finally, I used fluorescent multispectral beads to show that for samples where a high degree of colocalization between emitters in different spectral channels was expected, the final data processing pipeline recovered apparent labeling efficiencies of $\geq$93 % in experimental data.

The presented results demonstrate that gSEP in combination with `thunderSTORM` for emitter localization and `DOLanalysis` for colocalization analysis yields reliable labeling efficiency estimates across different imaging conditions and for labeling efficiencies between 5 and $>$90 %.

One important prerequisite to keep in mind is that the correction factors used to account for limited emitter detection efficiencies were determined with the help of simulated data. The simulations used in the present work to determine correction factors were performed such that emitter properties in simulated images matched the properties of emitters observed in the experimental data described in section 3.6. For future work it will be critical to ensure that experimental data and simulations employed for determining correction factors exhibit comparable emitter properties. This can either be achieved by ensuring that experimental data is acquired under identical conditions or by adapting simulations to match experimental emitter properties making use of the approach described in Section 3.4.1. As the applied degree of correction applied to experimental data depends on the detected emitter density, proper segmentation of cells in images and a homogeneous emitter density across the segmented area is a prerequisite which should always be verified for each set of experimental data. In the scope of this study, the segmentation quality and the homogeneity emitter densities in individual images was manually assessed.

If required, both points can be addressed to reduce the need for manual used input and data curation. Simulations used for obtaining correction factors can be automatically generated from acquired experimental data based on the properties of detected emitters in that data. The homogeneity of emitter densities across an image can also be assessed in an automated fashion to enable data curation without manual inspection of individual image sets by the user.

The optimized and validated pipeline for measuring labeling efficiencies was finally applied to study the influence of ligand concentration and incubation time, the influence of the cell line in which measurements were performed as well as the influence of fluorophores conjugated to ligands and the influence of chemical fixation on tag activity.

In comparison to the initial gSEP data analysis approach used by Haenselmann[174], labeling efficiencies determined for SNAP$_f$-tag and HaloTag in Huh7.5 cells across different incubation times and ligand concentrations after reanalysis of the same dataset exhibited a large degree of overlap in their qualitative behavior. However, labeling efficiencies for HaloTag labeling with HTL-TMR obtained after reanalysis of the dataset tended to be slightly lower than previously, and labeling efficiencies for SNAP$_f$-tag labeling with BG-SiR tended to be slightly higher than previously. The cell-to-cell variation at individual conditions was slightly reduced after reanalysis. As the observations made in Huh7.5 cells after reanalysis are in line with the relative behavior of SNAP$_f$-tag and HaloTag in new datasets measured in H838 and HeLa cells, it seems likely that the improved emitter detection approach resulted in a more consistent performance across varying conditions.

Across all different conditions I consistently observed that neither for SNAP$_f$-tag nor for HaloTag, stoichiometric labeling was achieved. This observation was in contrast to an initial report for labeling of SNAP$_f$-tag in vitro where labeling efficiencies ∼90 % were reported for different BG-ligands based on mass-spectrometry measurements**Sun:CorreaJr:2011**. The study presented here, provides, for the first time a systematic evaluation of achievable labeling efficiencies in living or fixed cells. Different groups reported labeling efficiencies measured using different approaches for individual labeling conditions which support the findings made in the course if this study.

Wilmes et al. used an assay based on single-particle tracking to determine the labeling efficiency for HaloTag labeling in live HeLa cells. After labeling with HTL-TMR at a ligand concentration of 30 nM for 15 min, a labeling efficiency of 22 % was determined[91]. Under comparable conditions, our experiments performed in HeLa cells indicate that saturated labeling should have occurred at the conditions used by Wilmes et al. a labeling efficiency of 26 % would be expected based on our measurements. Latty et al. made use of a colocalization-based approach with a fluorescently labeled antibody as reference marker to measure the labeling efficiency for HaloTag[189]. Here, labeling with HTL-TMR at a ligand concentration of 5 nM and labeling for 30 min in live HEK293 cells led to a labeling efficiency of 33 %. Under identical labeling conditions, but in H838 and Huh7.5 cells, I observed very similar labeling efficiencies of 38 ± 12 % and 36 ± 10 %.

For SNAP_f-tag, Virant et al. measured the labeling efficiency achieved upon labeling with BG-AF647 in cell lysates using a intensity-based readout with the FtnA complex as reference standard with defined stoichiometry[168]. Here, a labeling efficiency of 65 % was achieved for labeling with 100 nM ligand concentration for 6 h. Under comparable labeling conditions (fixed cells, 100 nM ligand concentration, 2 h incubation time), I observed a highly similar labeling efficiency of 69 ± 11 % (Section 3.6.5). A second study by Thevathasan et al. reported similar labeling efficiencies for SNAP_f-tag achieved with BG-AF647in an dSTORM-based assay using the nuclear pore complex as reference structure[92]. Here, labeling efficiencies of 56-64 % were determined under identical conditions as were used in my experiments with BG-AF647. For labeling of HaloTag with the ligand HTL-O2-AF647 at a concentration of 1 µM, Thevathasan et al. measured a labeling efficiency of 36 ± 3 % slightly less than the labeling efficiency observed by me with gSEP under the same labeling conditions (45 ± 1 %, Section 3.6.5). Using the same approach, Thevathasan et al. measured a saturation labeling efficiency of 21 ± 4 % for HaloTag with the photoactivatable ligand PA-JF549 across different ligand concentrations (0.25-5 µM). They reported constant labeling efficiencies for all tested ligand concentrations between 250 nM and 5 µM corresponding to the saturation labeling behavior observed in my data.

Taken together, these observations previously reported in the literature support my observations made with the gSEP approach and thereby further strengthen the claim that gSEP produces reliable estimates of labeling efficiencies achieved in situ and that labeling with those tags is far from stoichiometric.

In Section 3.6.2, I further showed that in addition to estimates of the saturation labeling efficiency, the labeling kinetics for a chosen combination of protein tag and ligand could be determined by modeling the labeling efficiency in dependence on incubation time and ligand concentration. Interestingly, the apparent rate constant for labeling of HaloTag with HTL-TMR determined in this fashion was about 10-fold lower as the value recently reported by Wilhelm et al. for measured with purified proteins in vitro[70]. This suggests that cell permeability can significantly influence labeling kinetics even for ligands such as HTL-TMR, which were previously reported to be highly cell permeable[74].

This is especially important since I observed a dependence of achieved labeling efficiencies on the fluorophore conjugated to ligands (Section 3.6.3). These experiments helped to distinguish between ligands with limited use in protein counting experiments due to their tendency for non-specific labeling (e.g. BG-JF549 and -JF646) and ligands such as HTL-JF646 which represent promising candidates for quantitative studies. In the future, systematic studies analogous to the experiments performed with BG-SiR and HTL-TMR in H838 cells will be useful to map out the labeling behavior and to identify optimal labeling conditions with maximum labeling efficiency and minimum non-specific labeling for alternative ligands.

One interesting finding highlighting the need for labeling efficiency measurements is the differential labeling efficiency achieved for SNAP$_f$-tag pre- and post fixation. An effect like this would have been easy to overlook if only labeling kinetics were determined. A comparison of labeling by comparing signal intensities pre- and post-fixation could easily be biased due to differences in the microenvironment experienced by fluorophores which are known to influence fluorophore properties (section 2.1).

The large scale comparison of labeling efficiencies across different tags, cell lines, incubation times and ligand concentrations showed that in its current implementation, the gSEP approach can be sensitive towards differences in the raw data used as input for emitter detection and subsequent colocalization analysis. As already mentioned in Section 3.6.1, deviations in input data were observed to lead to systematic deviations in the recovered labeling behavior of individual samples. While it is difficult, to determine the specific reason for deviating data quality, the variability in different properties of the acquired data was increased after switching from manual data acquisition to scripted, semi-automated data acquisition (Fig. A.4).

Improving the data acquisition routine with a focus on reproducibility across and within samples will therefore help to minimize the variability in measured labeling efficiencies, while retaining the ability to acquire large datasets with minimal user input. It is known that for TIRF illumination, the penetration depth of the evanescent wave and the intensity of signals depend strongly on the illumination angle[190]. Calibration of the illumination angle following different published approaches[190,191] prior to each experiment would therefore be expected to lead to better reproducibility in the obtained TIRF illumination and consequently reduce variations in signal amplitudes, as well as the SNR across samples. The homogeneity of TIRF illumination across the field of view is typically reduced by mismatches in the refractive index between different parts of the specimen[192]. To improve the evenness of TIRF illumination, azimuthal beam scanning on the back focal plane was developed to effectively reduce illumination inhomogeneities[193].If implemented with a motorized tip/tilt mirror, this approach comes at the additional advantage that it enables precise control of the illumination angle[192]. In the current study, a hardware-based autofocus based on total internal reflection of a near-infrared laser beam was used maintain the focal position during automated imaging of multiple cells in a sample. In practice, this system was found to not exhibit the stability required for fully automated imaging. Yasui et al. recently described a combined hardware and software autofocus system which employs a neural network for robust identification of the focal position and showed that the performance of this approach was sufficient for fully automated single-molecule fluorescence microscopy in living cells[194].

Complementary to achieving more robust data acquisition, several parameters obtained from emitter localization can be used for automated quality control of acquired data. As described in Section 3.6.1, the average detected emitter number in the reference channel across a dataset was found to be a predictor of samples with deviating properties. In addition, the amplitude and width of individual PSFs might be used on a per-cell or per-sample basis to automatically identify samples acquired under irregular acquisition conditions. Together, I expect the adoption of an imaging protocol and a more stringent quality control of input data to lead to a substantial improved robustness of the presented gSEP approach for labeling measurements based on single-molecule colocalization analysis.

In summary, the work presented in this chapter laid the foundation for routine labeling efficiency measurements for protein labeling with self-labeling protein tags. While future developments suggested above can help to further improve the robustness of the gSEP approach for high-throughput studies, the data acquisition and analysis pipeline in the current implementation are ready for use in routine applications. The labeling conditions identified in this chapter can further be used to achieve optimal performance for labeling of SNAP$_\mathrm{f}$-tag or HaloTag in emitter counting microscopy applications.

Chapter 4

IN SITU PROTEIN COUNTING WITH FLUORESCENCE MICROSCOPY

In the previous chapter, I presented a method for measuring the labeling efficiency upon protein labeling with genetically encoded tags. This enables relating measured fluorophore numbers to the underlying number of target proteins. With this tool at hand, it is possible to use fluorophore counting methods for molecular counting of target proteins. Photobleaching step analysis is a calibration-free fluorophore counting technique based on the identification of step-like intensity changes in intensity traces recorded by fluorescence microscopy during photobleaching of a sample. In this chapter, I will describe the development and validation of `quickPBSA`, a comprehensive data processing framework for fluorophore counting in situ via photobleaching step analysis. `quickPBSA` was developed together with Johan Hummert (Herten lab, University of Birmingham) who implemented the data analysis framework in Python and assisted with data acquisition and analysis. Significant parts of this chapter are part of a publication[195]. Johan Hummert and I are equal first authors of this publication. Theresa Fink contributed to the project in the scope of a lab rotation by performing initial measurements and helped with validation of `quickPBSA`.

4.1 Scope

As described in section 2.3, a number of PBSA approaches for interpretation of intensity traces recorded during photobleaching of samples have been developed in the past. However, for a PBSA method to be applicable for in situ fluorophore counting, i.e. in the complex environment of a cell, several criteria have to be met. First, the trace interpretation algorithm should not make implicit assumptions such as uniform photobleaching step height across a set of photobleaching traces, photostability or monotonous fluorescence decay without visible photo-induced blinking, which are rarely fulfilled for data acquired on cellular samples. Second, the complexity of eukaryotic cells will result in varying raw trace quality due to variable background signal, variable spacing between regions of interest and variations of axial mismatch between complexes and focal plane. The PBSA framework should thus be able to automatically extract photobleaching traces under such conditions and remove traces which do not pass a quality threshold. Third, many cell biological processes exhibit a significant degree of variability due to stochastically occurring events or heterogeneous populations reflecting different functional states. To capture and analyze this variability, a PBSA method for in situ analysis of protein complexes will have to be able to process a sufficiently large number of photobleaching traces on a computationally feasible timescale in an automated fashion and with minimal user input during analysis. Finally, any fluorophore counting technique will have to be fully validated using data with known ground truth. Here, validation using in-cell reference standards will always provide a more realistic picture of the performance of a new method compared to validation with simulated data or with experimental data acquired under idealized conditions.

Since we found that none of the currently available PBSA approaches was able to meet these criteria, we decided to develop `quickPBSA`, a dedicated framework for in situ PBSA building on previous approaches available in the field. This framework was intended to address three main shortcomings of available methods:

1. Automation: `quickPBSA` processes raw data and a list of candidate ROIs in an automated fashion.
2. Robustness: By implementing multiple filtering steps, we ensure robustness against variable input data.
3. Computation speed: In its current implementation, `quickPBSA` is suited for analysis of datasets containing 100s of traces on a simple desktop computer.

In the following, I will first describe the design principles of the `quickPBSA` framework for automated trace extraction and PBSA. I will then describe the development of the `quickPBSA` algorithm for trace interpretation. This will be followed by an in-depth validation of `quickPBSA` with two independent approaches using synthetic and experimental datasets with varying complexity. I will then show that `quickPBSA` performs robustly on data acquired on different microscopy setups, but that the choice of fluorophore can influence the results obtained with `quickPBSA`. Building on the rationale that the robustness of protein counting in situ with PBSA on high copy number targets depends on the quality of raw data, I performed a comparison of commonly used protein labels and imaging conditions with respect to their brightness, photostability and propensity for photo-induced blinking. Finally, I evaluated two previously described reference protein standards and characterized their performance as in-cell benchmarking targets. Using a genome-edited cell line expressing a NUP107-SNAP$_\text{f}$-tag fusion protein, I show that in combination with optimized data acquisition conditions, `quickPBSA` is able to measure Nucleoporin (NUP)107 copy numbers in individual eukaryotic nuclear pore complexes in situ.

4.2 Development of the quickPBSA framework

To allow for automated processing of PBSA data, we designed `quickPBSA` as a modular framework where each module serves a specific function and individual modules can, in principle, be replaced with alternative approaches. In cases where were already available for individual steps in our framework, we tried to make use of those and adapted them as needed for use as a module in the `quickPBSA` framework or established routines suitable for automated execution of the full data analysis pipeline. In the following, I will describe the core tasks executed in the individual `quickPBSA` modules and their respective implementation.

As shown in Fig. 4.1a, `quickPBSA` consists of several modules which are consecutively applied to a given set of raw timelapse images. We decided to split the trace interpretation into two separate modules to maximize the amount of information extracted from photobleaching traces, while at the same time minimizing the computational effort required to process a given dataset. The candidate ROI detection module is currently implemented as `ImageJ/Fiji` script, while all further modules are implemented in the `quickPBSA` python package[1]. The individual modules perform detection of candidate ROI positions, filtering of candidate ROIs, trace extraction, step detection, step filtering and model refinement, respectively. The following section will describe the function and underlying principle for each module.

Candidate ROI detection module: The first module in the `quickPBSA` framework serves to identify ROIs in the raw image data where photobleaching traces are to be extracted from. Different modes of ROI selection are supported by `quickPBSA`. ROIs with arbitrary

---

[1] Source code available via pip at https://pypi.org/project/quickpbsa/ or from github at https://github.com/JohnDieSchere/quickpbsa

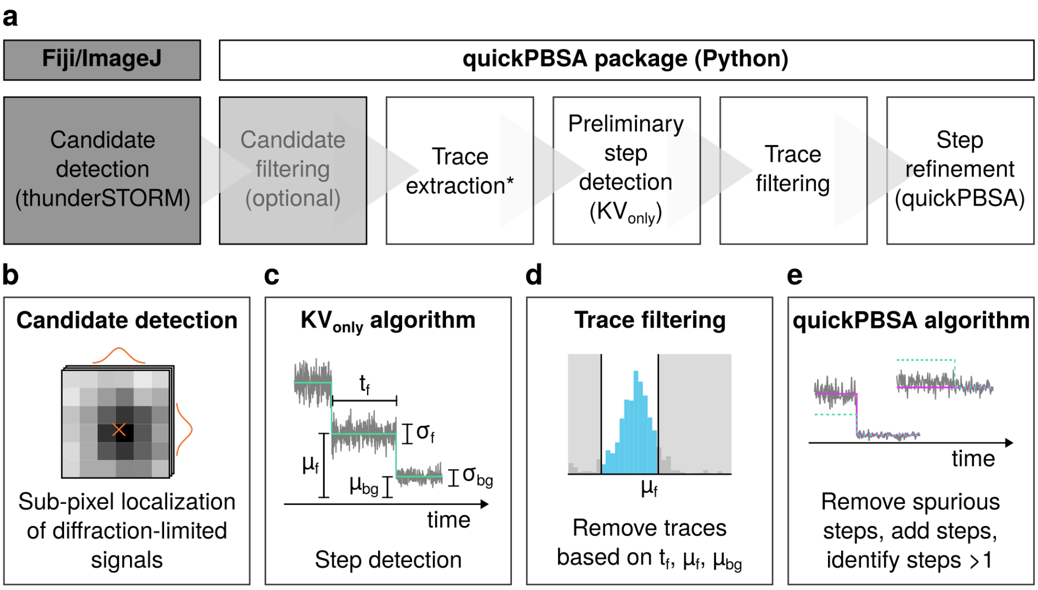

Figure 4.1 Illustration of quickPBSA workflow for PBSA. a, Initial localization of diffraction-limited structures is performed with external software for SMLM data analysis. All further steps can be performed using the `quickPBSA` package implemented in Python. Trace extraction can be performed with or without local background correction. See Fig. 4.27 for additional details about candidate filtering and trace extraction. **b-e** Schematic overview of detailed procedures performed during candidate detection, step detection (Kalafut-Visscher step detection (KV$_{\text{only}}$)), trace filtering and step refinement (`quickPBSA` algorithm). Trace models shown in green represent results from KV$_{\text{only}}$ algorithm, trace models shown in magenta represent results from `quickPBSA` algorithm.

shapes can be subjected to `quickPBSA` as binary masks created by hand or with thresholding or segmentation algorithms. For structures with diffraction-limited emission patterns, a second mode of ROI selection is available. Here, identification of diffraction-limited signals in raw images can be performed using any of the available software solutions for sub-pixel localization of emitters (Fig. 4.1b) in SMLM[175]. In the course of the current work, the `quickPBSA` framework was exclusively used to investigate protein clusters smaller than the diffraction limit. Emitter localization was performed using the `thunderSTORM` SMLM data analysis plugin for Fiji/ImageJ[178] (for additional details, see Section 6.3.5).

Candidate ROI filtering module: After candidate detection, the candidate filtering module allows for excluding candidate positions from downstream analysis based on predefined selection criteria. This module uses the output of the candidate detection module as input. If candidate ROIs were identified using the SMLM option, the list of emitter localizations and the properties of individual emitters can be used as additional selection criteria. Additional details regarding the importance of this module for in situ counting will be discussed in Section 4.5.3.

Trace extraction module: After identification and pre-selection of suited ROIs, intensity traces are extracted from raw image data at positions defined by the candidate ROI list. This extraction can be performed with our without local background correction using a ring-shaped background region around individual ROIs to obtain the background intensity. Additional details regarding background correction during trace extraction will be explained in Section 4.5.3. Once individual photobleaching traces have been extracted, `quickPBSA` proceeds with interpretation of these traces. If alternative approaches for ROI selection and/or trace extraction are to be combined with the `quickPBSA` trace interpretation module, the candidate ROI detection, filtering and trace extraction modules can be omitted from the pipeline.

Preliminary step detection module: Different methods for detecting step-wise changes in time series exist[196] and have been applied to photobleaching time traces. Here, a balance between sensitivity of event detection and overfitting is critical for successful identification of photobleaching steps while ignoring random intensity fluctuations. One algorithm, which was successfully applied for step detection in fluorescence intensity traces was developed by Kalafut and Visscher and is based on the Schwarz Information Criterion (SIC, see appendix on 224) as metric for deciding if and where step-wise intensity changes occurred[197]. The Kalafut-Visscher (KV) algorithm subsequently adds steps by computing the Schwarz infomation criterion (SIC) for all possible step positions and then accepting the position with minimal SIC value. This minimal SIC value of the trace model with k steps is then compared to the SIC value of the trace model without this step (k-1 steps). If the model with k steps exhibits a lower SIC value, the added step position is accepted. This procedure is repeated until the addition of an additional step does not result in an improved SIC.

In our implementation of the KV algorithm, which I will refer to as $\mathrm{KV}_{\mathrm{only}}$ in the subsequent sections of this chapter, we added three changes to reduce the computation time for preliminary step detection. First, we included a parameter *threshold*, which is used to exclude possible step positions with intensity differences smaller than *threshold*. This is to prevent minor intensity variations from being considered as steps and thereby increasing the computation time required during model refinement. Second, we introduced a trace cropping procedure to reduce trace lengths and thereby the number of possible step positions which need to be evaluated. If active, trace cropping discards data points starting from the end of a trace and moving towards the start of the trace until the difference between two data points exceeds 0.5**threshold*. To prevent trace cropping from interfering with the estimation of background intensity properties, additional data points specified by the parameter *bgframes* after the crop point are retained in the cropped trace. Finally, we added the parameter *maxiter* which specifies the maximum number of steps allowed to be placed during initial step detection to avoid spending excess computation time on traces which far exceed the expected complexity of traces. In practice, *maxiter* is set to values >5 times the expected maximum number of fluorophores to be contained in a complex.

The cropped trace in combination with the preliminary list of step positions is then used to compute key properties for each trace. The brightness of a single fluorophore μ_f and brightness variance σ_f of a single fluorophore are computed from all data points in the interval between the two last detected steps in a trace. The mean background intensity μ_{bg} and background variance (σ_{bg}) are computed from all data points after the last detected step in a trace. In addition, the length of the interval between the last two detected steps in each trace is extracted (see Fig. 4.1c). The output from this module thus consists of a sequence of potential step positions as well as a set of properties for each trace.

Trace filtering module: After initial step detection using the KV_{only} algorithm, the obtained preliminary step positions are used to decide which traces to include in downstream analysis. For this, the distribution of single fluorophore brightness (μ_f) is computed across all traces in a dataset. Two user-defined parameters are used during trace filtering: *percentile_step*, is a cutoff to exclude traces where μ_f is larger than the percentile of the distribution of all μ_f across a dataset. *length_laststep* is the minimum number of data points between the last and the second-to last detected step. Traces are excluded from further analysis based on a list of requirements, which each render successful processing in the model refinement module unlikely or impossible:

- KV_{only} algorithm detected no steps indicating that either, no step-like changes in intensity occurred or these changes were below the pre-defined intensity *threshold* which could for example be the case for traces extracted from structures located far away from the focal plane. In both cases, a meaningful estimate of fluorophore numbers initially present in a ROI is not possible.

- Parts of intensity trace are below -*threshold* after local background correction. If background detection does not perform as expected, e.g. due to missed adjacent fluorescent structures during candidate ROI detection, the background correction routine will overestimate the background signal which may lead to traces exhibiting intensities below zero.

- KV_{only} detected more positive than negative steps. Such behavior is typically observed if the background correction routine overestimates local background fluorescence.

- μ_{bg} is out of bounds defined by [-*threshold*,*threshold*] for traces after correcting for local background. If the intensity after the last detected step is substantially higher or lower than expected based on the specified minimum brightness of a single fluorophore (*threshold*), this is an indicator that not all fluorophores were completely photobleached at the end of the measurement or that local background subtraction did not work as expected.

- Single fluorophore brightness μ_f is out of bounds [*threshold*,*percentile_step*] after local background correction. If the height of the last detected step in a given trace is lower than *threshold* after local background subtraction, this can be an indicator for a structure within the ROI being located outside the focal plane or that the ROI used for trace extraction was placed in a way that it only partially captured signal emitted from the structure. If the height of the last detected step in a given trace is higher than *percentile_step*, this can be an indicator for more than a single fluorophore being photobleached within the last detected step. To exclude both cases, only traces with μ_f in the given range are accepted (Fig. 4.1d).

- The interval between the final two steps detected by KV_{only} is shorter than *length_-laststep*. Although such traces, in principle, could be suited for downstream refinement, a short interval between the last two steps will limit the quality of μ_f and σ_f estimates and thereby the reliability with which these values can be used during refinement.

By including this module between the preliminary step detection and model refinement modules, we avoid unnecessary evaluation of traces in the model refinement module, which should help to minimize the computation time required for processing a given set of traces.

Model refinement module: The KV_{only} algorithm proved to be highly sensitive for detecting step-like changes within a trace, but lacks the ability to distinguish steps likely arising due to fluorophore photobleaching from signal fluctuations. KV_{only} is also not able to distinguish between steps where a single and multiple fluorophore bleaching events events occurred simultaneously. As explained above, the characteristics of a single fluorophore and the local background can be obtained from a given trace. For this reason, we decided to make use of these information in an additional module. This model refinement module re-evaluates the predicted sequence of photobleaching steps (i.e. the trace model) obtained from preliminary step detection in the light of fluorophore and background characteristics.

This concept builds on a similar idea recently presented by Tsekouras and co-workers, who used a Bayesian framework for interpreting photobleaching traces[123].

In their approach, Tsekouras et al. define a posterior, which takes the reduced probability for simultaneous photobleaching events where more than one fluorophore photobleaches and the higher likelihood for photobleaching events at the beginning of a trace where more fluorophore are still present into account. This posterior is then used to compare a set of different models and to identify the most likely model as series of photobleaching and photoblinking events which best describes the photobleaching trace. In practice, this selection is made by computing the negative logarithm of the probability $P(\theta|D)$ given an intensity trace D and parameter set θ (see Equation B.3 on page 224 for definition of posterior and corresponding parameters in Table B.1).

Since a systematic comparison of every possible model is computationally not feasible, Tsekouras et al. apply a model pre-selection strategy to discard unlikely models from further evaluation. This selection strategy divides the photobleaching trace into discrete windows for which a likely range of active fluorophores is determined based on single fluorophore and background intensity and variance determined using the Kalafut-Visscher algorithm (KV) algorithm. For each window, the most likely active sub-model is then identified using Equation B.3. In this approach, both the model pre-selection and precise model selection step contain an additional adjustable parameter for restricting the parameter search ranges which each influence both, computation time and accuracy of the algorithm. The algorithm by Tsekouras et al. performed well on simulated photobleaching traces with up to 50 fluorophores present at the beginning of the trace and required shorter computation time compared to previous approaches (17 h vs. ∼168 h for a trace with 50 fluorophores and 5×10^4 data points[115]).

For the `quickPBSA` model refinement module, we rely on the same posterior as used by Tsekouras et al., but apply a different strategy for pre-selection of likely models, which are compared using the posterior. Instead of choosing a windowed approach to limit the number of tested models, we reasoned that, firstly it is unlikely that a photobleaching step occurs for most positions along the trace and, secondly that the step positions detected by the KV_{only} algorithm provide a reasonable starting point for subsequent refinement of the photobleaching event sequence. As a result, the search space for the best model to describe a given photobleaching trace is massively reduced since the size of the parameter space is mainly determined by the number of data points, i.e. the number of possible photobleaching event sites in a trace. As an additional strategy to restrict the search space during model refinement, we follow Tsekouras et al. and use the single-fluorophore and background characteristics determined in the preliminary step detection module for individual traces as constants during model refinement. To further minimize the computation cost of the refinement process, we devised a scheme for iterative evaluation of subsets of likely models without systematically testing all possible combinations of altered step properties and positions.

The refinement procedure is initialized with the step position sequence determined in the preliminary step detection module with the $\mathrm{KV}_{\mathrm{only}}$ algorithm. All models are evaluated by computing $-2\ln P(\theta|D)$ according to Equation B.3. The model exploration scheme consists of four steps, which are iteratively applied to each trace and the corresponding preliminary photobleaching event sequence as illustrated in Fig. 4.2a and described in the following:

Initialize by computing posterior according to Equation B.3 for the step position sequence determined by the $\mathrm{KV}_{\mathrm{only}}$ algorithm.

(I) Generate list of candidate positions for occurrence of steps with heights >1. For the first iteration, consider all positions as candidate positions. For later iterations, only consider positions where previous increase in step height improved model. If candidate positions are identified, proceed with II. If no candidate positions are left, proceed with III.

(II) Generate a set of models with all possible combinations of candidate positions and step heights up to the current iteration. For example, in iteration 3, step heights of 1-3 are considered. Keep all positions where increase in step height led to improved model as candidate positions and return to I.

(III) Continue with best model from I-II and remove step position with lowest SIC value from preliminary step detection. Then reset all step heights and return to I. Always keep the last two steps in a trace since these steps are required to calculate the posterior. Once only three step positions are left, proceed with IV if best current model contains the same number of step positions as the initial model from preliminary step detection.

(IV) Create and evaluate additional models by adding positive or negative steps at each time point before the second-to-last preliminary step position. Continue adding steps until no further improvement in -log(P) is observed or the maximum number of allowed steps has been added.

From all models created during I-IV, choose model with minimal -log(P) as final model. The effect of model refinement in removing spurious steps, adding additional photobleaching steps and adjusting the number of fluorophore photobleached in a step is illustrated in Fig. 4.2b.The number of possible models, especially for traces where many possible step positions were detected during preliminary step detection, can become computationally expensive. We added break conditions to each of the refinements steps described above to limit the evaluation of increasingly unlikely models and thereby reduce computational cost. A summary of parameters in the `quickPBSA` framework is given in Table B.2. Additional information on in- and output data for `quickPBSA` and parameters for optimizing computational performance can be found in the `quickPBSA` documentation[1]. Typical settings used for data analysis in the course of this chapter are listed in Section 6.3.5.

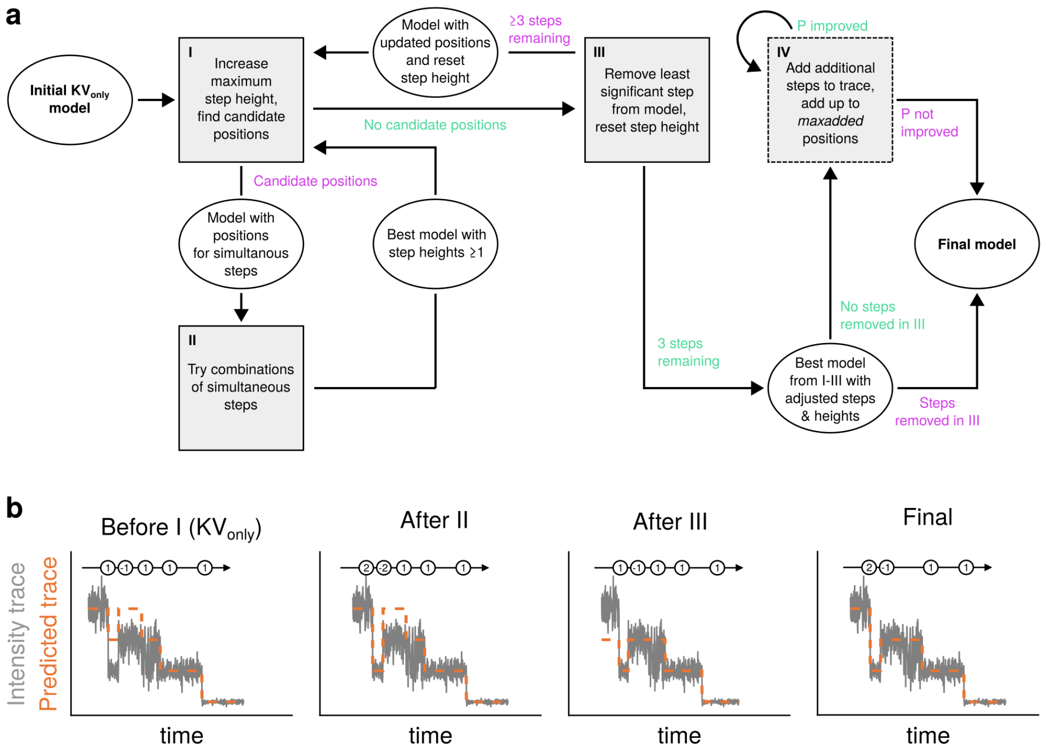

Figure 4.2 Illustration of the quickPBSA framework and algorithm. a, Schematic description of `quickPBSA` model refinement algorithm describing operations used for refinement based on result from preliminary step detection using $\mathrm{KV}_{\mathrm{only}}$ algorithm. Boxes describe operations, which modify step positions or properties. Ovals: Model states. Customizable parameters important for individual operations italicized. See Table B.2 for explanation of individual variable. **b,** Comparison of models (orange) and raw data (gray) after different model refinement steps. In the case of the trace shown, a step was removed in step III and step IV was therefore skipped. Data in b) was obtained from Johan Hummert.

---

[1] https://github.com/JohnDieSchere/quickpbsa

Local intensity algorithm: The signal recorded from a ROI at t_0, i.e. the beginning of data acquisition, depends on the number of active fluorophores and the brightness of individual fluorophores. For benchmarking of the `quickPBSA` framework, we decided to implement an alternative way to estimate the fluorophore number at t_0 in a given trace based on calculating the ratio between intensity at t_0 and single-fluorophore brightness. Since the observed brightness of fluorophores depends on multiple factors, we reasoned that using a global or even external calibration of single fluorophore brightness, as has been suggested by others[124], to likely be unreliable. Therefore, we decided to use single-fluorophore brightness μ_f determined in the preliminary step detection module as internal standard for the respective trace. Such an approach has the advantage of not relying on correct interpretation of complex traces and only requiring identification of the final two steps in a trace which is usually less demanding on the step detection sensitivity. The precision of such an approach, however, depends strongly on the SNR and fluorophore photophysics. If the SNR is low or if fluorophores frequently reside in non-emissive states, a large uncertainty of the t_0 fluorophore estimation is expected. In the course of this chapter, I will refer to this approach as local intensity approach.

4.3 Validation of quickPBSA with semi-synthetic data

In the previous section, I described the design of the `quickPBSA` framework for PBSA with individual modules for trace extraction, preliminary step detection, trace filtering and finally model refinement to determine the number of fluorescent emitters contained in microscopic structures. Before testing the performance of `quickPBSA` on experimental data, we performed a benchmarking with simulated data to validate the step detection and model refinement modules with data for which the ground truth is known. This initial set of experiments was also intended to compare the performance of `quickPBSA` with the previously reported approaches by Kalafut and Visscher[197] and Tsekouras et al.[197]. Typically, validation of PBSA algorithms is performed with fully synthetic data[124]. This has the advantages that parameters such as the SNR, the photostability and the photophysics of emitters contributing to signal are fully controllable. However, such an approach will, by design, only include known sources of signal variations and therefore likely underestimate the complexity of experimental data. For this reason, we generated semi-synthetic intensity traces from experimental data for which the ground truth number of active emitters could be determined by visual inspection.

To generate semi-synthetic intensity traces, I used DNA oligonucleotides probes labeled with four ATTO647N fluorophores per oligonucleotide (tetraATTO647N). The average fluorophore number per tetraATTO647N was determined previously using CoPS and was found to be close to 4[198]. tetraATTO647N probes were sparsely immobilized on the surface of cleaned glass coverslips via biotin-streptavidin linkage as illustrated in Fig. 4.3a. Immobilized probes were then imaged under high-intensity illumination ($\sim$2.4 $\mathrm{kW\,cm^{-2}}$) in photostabilizing ROXS $NaSO_3$ buffer (see Methods). Immobilized probes could be identified as individual diffraction-limited spots, which exhibited a decrease in number and intensity over time (Fig. 4.3b). While bleaching of individual spots was stochastic, the decay of intensity integrated from all spots in an acquisition was highly reproducible between acquisitions (Fig. 4.3c). Intensity traces containing signals from tetraATTO647N probes with up to four fluorophores per probe were then extracted from time-lapse image stacks by summing across the pre-selected ROIs around individual diffraction-limited spots. The extracted traces show the expected step-wise decrease in signal over time with step numbers varying between one and four (Fig. 4.3d). A high SNR and relatively few photoblinking events allowed us to determine the number and position of photobleaching steps that occurred in each trace by visual inspection.

To evaluate the performance of `quickPBSA` with respect to the maximum number of photobleaching steps per trace that could be robustly identified, we proceeded by constructing semi-synthetic traces by combining several low complexity traces according to the workflow shown in (Fig. 4.4a). First, manually annotated traces were filtered to select traces from probes, which were imaged with comparable illumination intensities[1]. The set of traces obtained after filtering showed a rich set of different features including transient on-off blinking, and transient states with lower brightness or complex intensity fluctuations, which are typically not encountered in simulated data (Fig. 4.4b, see Fig. B.1 for overview of all annotated traces). Complex traces with many photobleaching steps per trace were then generated by frame-wise summation of a random subset of traces (Fig. 4.4c). By altering the number of combined traces, the number of photobleaching steps in a traces could be varied. In total, this procedure was used to generate a dataset with $>4 \times 10^3$ semi-synthetic traces containing 2-41 active fluorophores at the beginning of each trace (Fig. 4.4d).

The generated semi-synthetic dataset was then used for evaluating the performance of the `quickPBSA` step detection and model refinement modules. For this, we first processed the entire dataset and then visually compared the `quickPBSA` trace models with the underlying

---

[1] Due to the Gaussian illumination profile of the laser used as light source, the illumination intensities across the field of view varied. See Section 5.2.1.

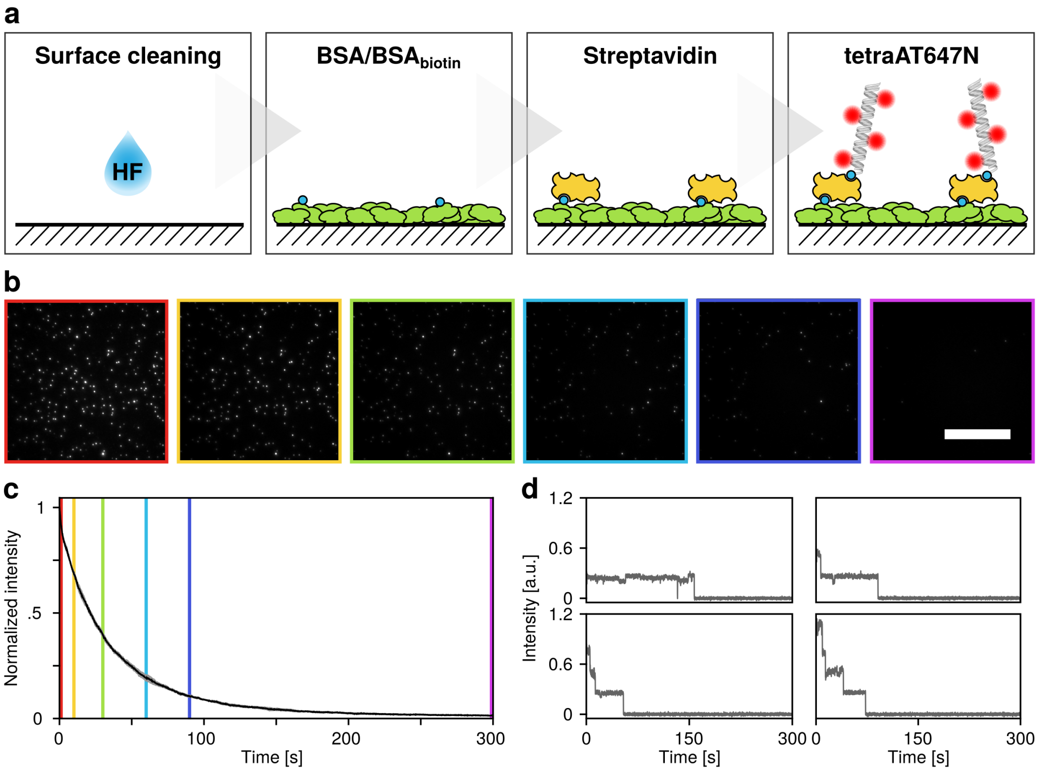

Figure 4.3 Acquisition of photobleaching traces from immobilized tetraAT647N probes. a, Workflow for immobilizing biotinylated DNA probes. Glass coverslips were cleaned with 0.1 M hydrofluoric acid (HF) and consecutively incubated with BSA/BSA$_{\text{biotin}}$, streptavidin and tetraAT647N (see Section 6.1.6). **b,** Individual frames from bleaching sequence at time points indicated by colored lines in c. Images displayed with identical lookup table settings. Scale bar: 20 µm. **c,** Photobleaching traces obtained by frame-wise summation of intensities within ROIs around bright spots in first frame. Mean (black line) ±SD (shaded region) from 9 FOVs. Please not that due to the high degree of similarity between measurements, the SD in no point is larger than 5 % therefore is barely visible. Traces were normalized to intensity in first frame. **d,** Representative photobleaching traces exhibiting 1-4 photobleaching steps extracted from individual diffraction-limited areas using `quickPBSA`. See Fig. B.1 for complete overview of traces extracted from dataset. Samples were imaged via widefield epifluorescence microscopy in ROXS NaSO$_3$ buffer. Image sequences were acquired at 1.3 kW cm^{-2} and 50 ms exposure time per frame until >90 % of spots were fully bleached ($\sim 6 \times 10^3$ frames).

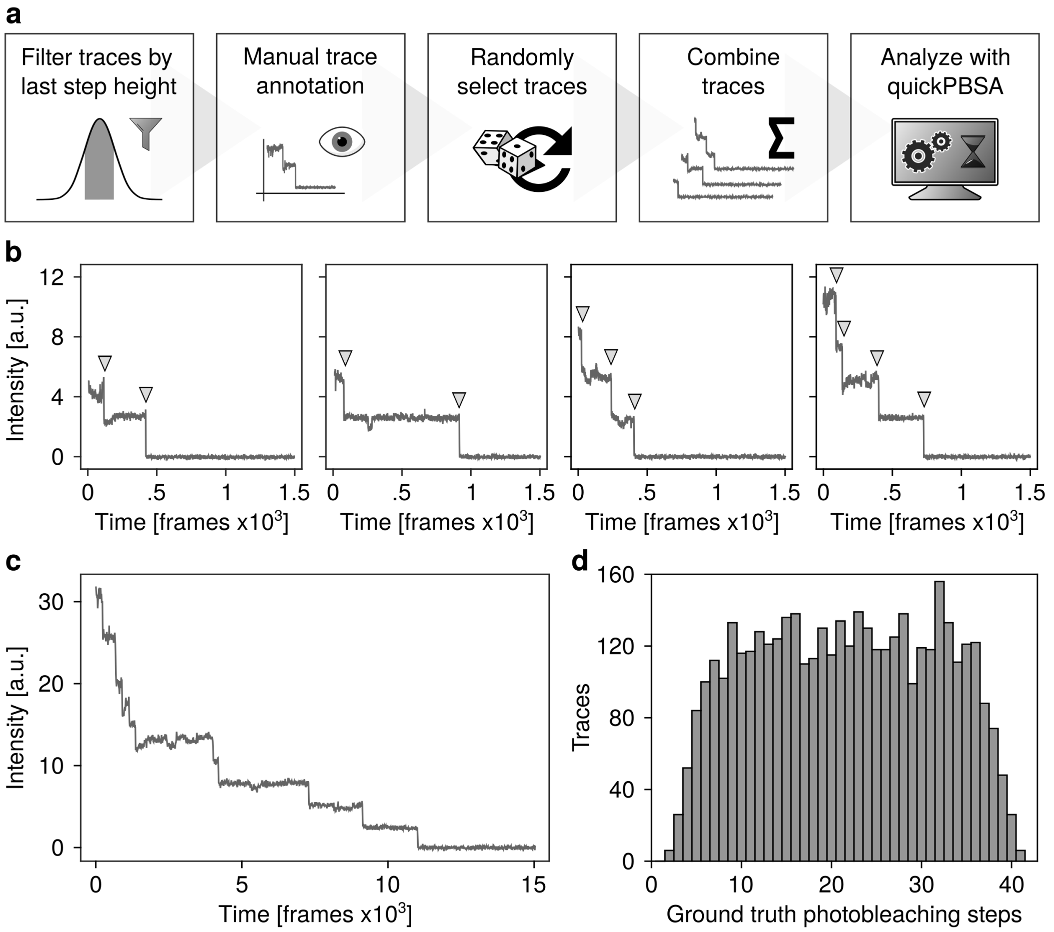

Figure 4.4 Construction of complex semi-synthetic traces from manually annotated photobleaching traces. **a,** Scheme depicting workflow for construction of semi-synthetic traces. **b,** Representative base traces selected for annotation with annotated photobleaching events indicated by arrowheads. **c,** Trace obtained by summation of base traces shown in b. **d,** Number of constructed semi-synthetic traces with respective total number of photobleaching steps in base traces used for construction.

input traces after linearly scaling the models produced by `quickPBSA` the absolute intensities of raw data. This scaling procedure facilitates a qualitative comparison between raw data and model. As shown in Fig. 4.5a, `quickPBSA` correctly interpreted different features of measured data such as short- and intermediate timescale blinking as well as minor intensity fluctuations not caused by photobleaching and finally predicted the correct number of fluorophores present at the beginning of the measurement. Events where more than one fluorophore had photobleached during a single frame were also correctly identified by `quickPBSA` (4.5b). Manual inspection of trace models returned by `quickPBSA` for raw traces with 15 and 30 ground truth photobleaching events further showed that `quickPBSA` is able to identify the correct number of photobleaching events even for traces containing a high number of events (4.5b,c). To systematically compare the performance of `quickPBSA` in a quantitative fashion, we compared the number of active fluorophores at t_0 predicted by `quickPBSA` with the known ground truth for the full set of $>4 \times 10^3$ semi-synthetic traces. Overall, `quickPBSA`was able to recover the expected number of events with high precision across the range of simulated photobleaching numbers (Fig. 4.5d). The deviation of the mean predicted fluorophore number for a given ground-truth fluorophore number was <1 across all fluorophore numbers tested (2-41 fluorophores at t_0). However, we also observed that `quickPBSA` showed a weak tendency towards overestimating the number of photobleaching events. This effect tended to increase slightly for intermediate fluorophore numbers (>15) before decreasing again for ground truth fluorophore numbers >30 (Fig. 4.5d).

To further assess the performance of `quickPBSA` with increasing emitter numbers, we compared the spread of `quickPBSA` fluorophore number estimates for increasing ground truth fluorophore numbers. Here, we observed that only a very weak broadening of the distribution of fluorophore number estimates for a given number of ground truth fluorophores occurred (Fig. 4.5e). In contrast, fluorophore number estimates across the set of semi-synthetic traces with the KV_{only} algorithm showed a stronger overall tendency towards overestimating the number of fluorophores contributing to a trace. This effect was correlated with increasing number of ground truth fluorophores (Fig. 4.5f). The observation, that `quickPBSA` provided better emitter number estimates in comparison to the KV_{only} algorithm shows that the `quickPBSA` model refinement module was able to correct the initial estimates from the KV_{only} algorithm, which is run during the preliminary step detection. The local intensity algorithm showed a tendency towards undercounting the true number of ground truth fluorophores (Fig. 4.5f).

Finally, we examined if the processing time of traces with `quickPBSA` depends on the number of fluorophores present at t_0 and hence the number of steps occurring in a trace and observed a weak dependence of the processing time on the number of ground truth fluorophores (Fig. 4.5g). For up to ~30 fluorophores, the computation was dominated by the preliminary step detection module which relies on the KV_{only} algorithm. For fluorophore numbers >30, the model refinement module based on the `quickPBSA` algorithm starts to contribute to the overall processing time in a significant manner. However, even for traces containing >40 fluorophores, the average processing time per trace was still below 10 s on a standard desktop computer (see Section 6.3.5 for specifications). In this respect, it is important to note that all semi-synthetic traces contained in this dataset had the same number of data points per trace (3×10^3). We expect that the number of data points per trace will influence processing time.

Having shown that `quickPBSA` is suited for automated PBSA, we compared `quickPBSA` to the related approach by Tsekouras et al.[123]. Here, our first goal was to determine whether differences between both approaches with respect to their precision or computational cost could be observed. The comparison between both approaches was performed using a subset of 40 traces from the semi-synthetic dataset described above with 2-8 ground truth photobleaching steps per trace and 3×10^3 data points per trace each. The algorithm developed by Tsekouras et al. was used in an implementation provided by the authors [1].

In the evaluated range of ground-truth photobleaching steps, both algorithms were able to precisely determine the number of expected steps per trace and no trend toward over- or undercounting was observed for either of the two algorithms (Fig. 4.6a). The overall error in step identification, defined sum of missed and excess steps over total number of ground truth steps, was about 5-6 % for both, `quickPBSA` and the approach by Tsekouras et al.. When comparing both approaches based on their computational cost, substantial differences were observed. For `quickPBSA` the total processing time including preliminary step detection and trace filtering was on average 4.9 ± 0.1 s (mean±SD) per trace, while the Bayesian approach by Tsekouras et al. required an average computation time of 76 ± 16 min (mean±SD) per trace (Fig. 4.6b). This means, that the optimized model refinement step developed in the `quickPBSA` framework resulted in an >900x reduced analysis time for the analyzed set of semi-synthetic traces. Interestingly, no systematic dependence of the runtime on the number of steps contained in a trace was observed for `quickPBSA` or the approach by Tsekouras et al. across the tested range of photobleaching steps. One important difference between the

---

[1] Download via https://github.com/lavrys/Photobleach

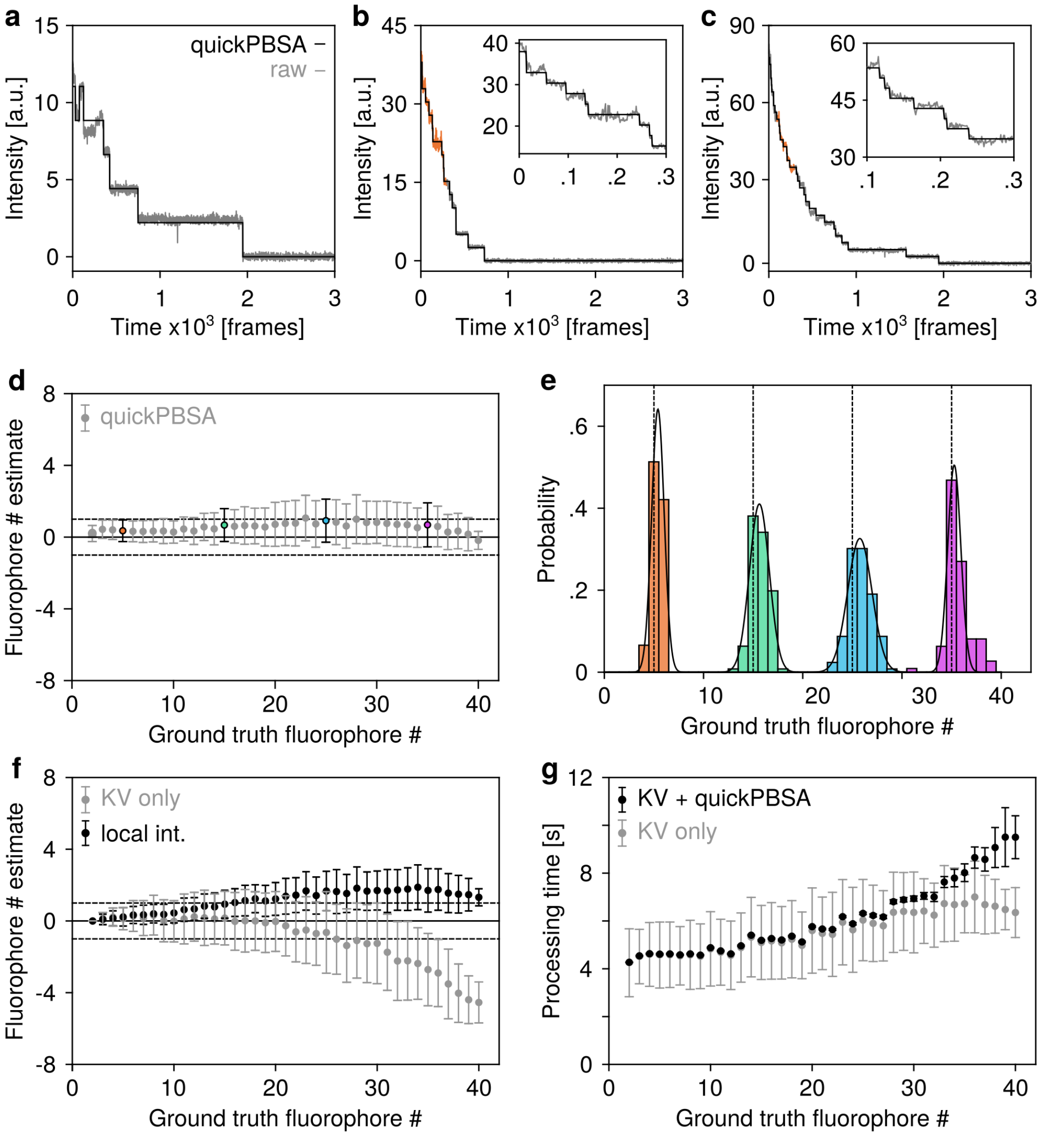

Figure 4.5 Photobleaching trace analysis of semi-synthetic data with quickPBSA. Semi-synthetic traces constructed using data recorded on tetraAT647N samples were processed using the `quickPBSA` framework. **a-c,** Raw intensity traces (gray) and corresponding `quickPBSA` emitter number estimates (black) for successfully analyzed traces with 5 (a), 15 (b) and 30 (c) bleaching steps. Emitter numbers were overlayed with intensity traces using a scaling factor (Section 6.3.5). b,c Insets: Zoom-ins to trace intervals highlighted in orange. **d** Fluorophore number estimates using `quickPBSA` for varying ground truth fluorophore numbers. Mean±SD. Highlighted values in e shown in matching colors. Black line indicates ground truth, dashed lines indicate deviation of ±1 from ground truth. **e,** `quickPBSA` emitter number estimate distribution (colored) for selected ground-truth bleaching steps (dashed lines). Distributions were modeled with a Gaussian distribution (black lines). **f** Fluorophore number estimates using $\mathrm{KV_{only}}$ (gray) and local intensity (black) approaches for varying ground truth fluorophore numbers. Lines as in (d). **g,** Processing time in dependence of ground truth fluorophore number. Mean±SD for processing with complete `quickPBSA` framework (black) and KV only step (gray). Data produced jointly with Johan Hummert.

approach by Tsekouras et al. and `quickPBSA` is that in the `quickPBSA` framework employs a filtering step to remove traces, which are unsuited for further interpretation in the model refinement module. To assess to which degree filtering of traces impacted the overall analysis, we determined the fraction of successfully analyzed traces at varying trace complexity in the semi-synthetic dataset described above. This analysis revealed that `quickPBSA` was able to successfully process $>83\,\%$ of all traces across the full range of ground truth photobleaching step numbers independent of trace complexity (Fig. 4.6c).

Taken together, these data show that `quickPBSA` is able to robustly estimate the number of active fluorophores in semi-synthetic photobleaching traces with up to $\sim$41 fluorophores. Furthermore, it was confirmed that the processing time required for trace analysis is sufficiently small to allow for processing datasets containing $\gg 1 \times 10^3$ traces without the requirement for specialized computation infrastructure, which represents a significant advantages over prior approaches.

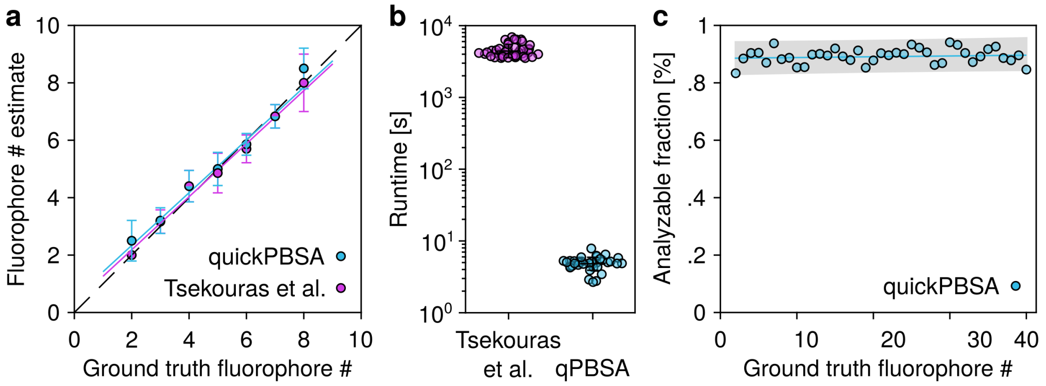

Figure 4.6 Comparison of quickPBSA with a Bayesian photobleaching step detection algorithm. The performance of both algorithms was compared using a set of 40 semi-synthetic photobleaching traces with 2-8 photobleaching steps and 3×10^3 datapoints per trace. **a,** Fluorophore number estimates (mean±SD)for `quickPBSA` (cyan) and Tsekouras et al.[123] (magenta) with linear fits to mean fluorophore number estimates (colored lines). Dashed line: expected number of detected photobleaching steps. **b,** Comparison of analysis runtime. For both algorithms, the runtime for complete analysis of individual traces (circles) was determined. Mean±SD indicated by error bars. **c,** Fraction of traces successfully evaluated using `quickPBSA` for varying photobleaching step numbers. Blue line: Linear fit with slope 3.7×10^{-4} and R^2 0.03. Shaded region: $95\,\%$ confidence interval of linear fit. The complete set of semi-synthetic traces described in Fig. 4.4 was used for analysis in (c).

4.4 In vitro emitter counting with quickPBSA

In the next step, I sought to confirm the findings made above in an experimental setting using structures carrying defined fluorophore numbers as targets. I reasoned that in order to achieve optimal performance for emitter counting with `quickPBSA`, optimizing the data acquisition conditions in a way that enabled recording of photobleaching traces with maximum SNR. In this section, I will therefore first describe a set of experiments performed to optimize data acquisition conditions with small organic fluorophores in vitro. I will then present experiments performed to characterize DNA origami which were used as target with known stoichiometry. In the last part of this section, I will evaluate the performance of `quickPBSA` in and previously used algorithms on DNA origami with varying stoichiometry and two different fluorophores.

4.4.1 Optimizing data acquisition with ATTO647N

The ability of PBSA algorithms to detect step-like intensity decreases in intensity traces depends on different factors related to both, the algorithm used and the input data that is being processed. Regarding the parameters of data acquisition, it is clear that higher SNR, simplifies the distinction of individual steps. Following this assumption, acquisition at the maximum possible excitation intensity should result in the best possible intensity traces for subsequent PBSA analysis. However, this would cause fluorophores to photobleach too rapidly to identify individual bleaching steps. Therefore it is required to find experimental conditions which provide high enough SNR but avoid rapid photobleaching of all fluorophores in short time.

An experimental strategy to address these contrary demands, is to choose bright and photostable fluorophores and to acquire data under conditions, which maximize the photon budget of fluorophores before they irreversibly photobleach. As explained earlier, this has been achieved for small organic fluorophores by fluorescence imaging in photostabilizing buffer systems, which remove molecular oxygen and which contain reducing and oxidizing agents that promote the transition of fluorophores from metastable dark states into the ground state (section 2.1). For in vitro validation of `quickPBSA`, I therefore chose the bright and photostable organic fluorophore ATTO647N as label for DNA origami in subsequent experiments. To maximize the photon budget of ATTO647N, I compared different photostabilizing buffer systems to identify an acquisition buffer with the highest photostability for this fluorophore.

To compare the photostability of ATTO647N conjugated to DNA in different buffers, I again made use of tetraATTO647N probes and immobilized them as described in the previous section (Fig. 4.3). I then chose to compare three photostabilizing buffers, which differ with respect to the system used to deplete molecular oxygen from solution. For the first buffer system, I used the well-established enzyme combination glucose oxidase (God) and catalase (Cat), which consumes molecular oxygen and glucose in a two-step reaction[199]. The GodCat system was found to effectively reduce the concentration of molecular oxygen dissolved in solution[27]. However, it was also shown that this system is prone to causing a decrease in pH during prolonged measurement times[28] as well as a difference in the enzymatic activity of glucose oxidase and catalase which might result in buildup of hydrogen peroxide. Therefore, I included an alternative enzymatic oxygen depletion system, which is based on the conversion of protocatechuic acid (PCA) by Protocatechuate-3,4-dioxygenase (PCD) to β-carboxy-cis cis-muconic acid under consumption of oxygen[200], as second buffer into my comparison. The PCD oxygen scavenging system was shown to achieve lower dissolved oxygen concentrations compared to the GodCat system[27]. The third buffer was based on a recent report which made use of sodium sulfite as oxygen scavenger[201]. Sodium sulfite $NaSO_3$ reacts with molecular oxygen to form sodium sulfate and was found to be able to efficiently deplete molecular oxygen from aqueous solution without change in pH and without the need for expensive enzymes[201,202]. To all three buffers, I added methyl viologen and ascorbic acid as ROXS components following typical buffer formulations for single-molecule fluorescence microscopy[32].

To compare the different buffer systems, I acquired time-lapse images in samples with all three buffer systems as well as in Phosphate buffered saline (PBS) as reference. I then extracted intensity decays by creating a global mask containing the PSFs of all visible tetraATTO647N in a FOV and extracting the integrated frame-wise intensity. For all four samples, I observed a constant decay of overall intensity over time with small variations between FOVs and experiments (Fig. 4.7a). The observed decays required fitting with a bi-exponential function indicating a more complex photo-induced reaction system with multiple processes occurring in parallel (Fig. B.3). Such behavior was also previously observed for different fluorophores and imaging conditions[203,204]. To compare buffer systems based on a model-independent metric, I determined $t_{1/2}$, i.e. the time at which the overall intensity had decayed by 50 % and normalized $t_{1/2}$ to an excitation power density of $1\,\mathrm{kW\,cm^{-2}}$.

While rapid bleaching within seconds occurred for samples imaged in PBS or the photostabilizing buffer based on $NaSO_3$. Addition of ROXS reagents and $NaSO_3$ approximately

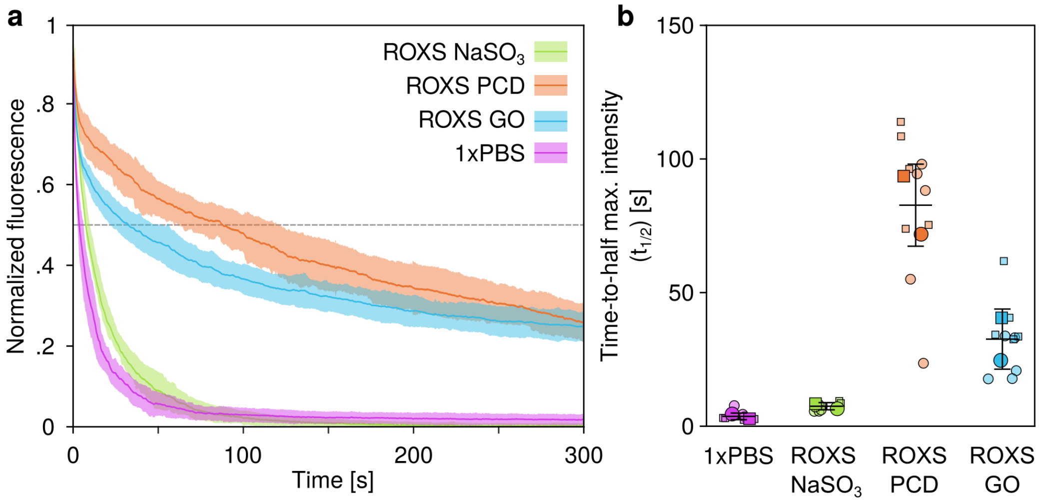

Figure 4.7 ATTO647N photostability in ROXS buffers. tetraATTO647N probes were immobilized on chambered glass coverslips via biotin-streptavidin linkage at single-molecule density (see 4.3a for representative images) and imaged in indicated buffers. The intensity over time was recorded from ROIs obtained by applying an intensity threshold. 640 nm excitation light at 1.26 kW cm^{-2} average power density was used for imaging. For each buffer system, a minimum of nine FOVs from two independent experiments were recorded. **a,** Intensity decays from each FOV were normalized to the intensity at time zero and averaged over all FOVs (lines). The standard deviation across FOVs for each condition is shown as shaded region. Dashed line: $t_{1/2}$. **b,** Time to half maximum intensity ($t_{1/2}$) corrected for excitation power for individual FOVs (small symbols) and means per experiment (large symbols). Error bars indicate mean±SD across experiments.

doubled the $t_{1/2}$ compared to PBS. Considerably slower photobleaching was observed in buffers with enzymatic oxygen removal, where a 10- to 30-fold improvement of $t_{1/2}$ for ROXS supplemented with GodCat and PCD systems, respectively was observed (Fig. 4.7b). These data suggest that changing the buffer system used for data acquisition from standard physiological buffers such as PBS to a photostabilizing buffer with enzymatic oxygen removal and a ROXS system should enable recording raw data with substantially improved properties for subsequent PBSA.

4.4.2 Characterization of DNA origami standards

Due to their defined properties and high stability, DNA origami represent ideal validation targets for quantitative fluorescence microscopy[156]. Different types of origami are available for benchmarking microscopes and acquisition schemes with respect to their achieved resolution and based on the brightness of detected objects[109]. In contrast to other biological macromolecules or macromolecular assemblies, DNA origami are highly modular and can be designed to exhibit different properties such as different physical size or different numbers of conjugated biomolecules[163]. Their modularity in combination with their low complexity makes them ideal standard targets with low sample-to-sample variation.

In this study, I used three different types of commercially available DNA origami. All three origami types had the same rectangular layout with sub-diffraction size (~70x90 nm) and were functionalized with a varying number of small organic fluorophores and biotin (Fig. 4.3a). Fluorophores are typically attached to the DNA origami base through binding of short complementary DNA oligonucleotides, which each carry individual fluorophores. By varying the number of complementary oligonucleotides used for labeling the origami base, the number of fluorophores attached to the base can be controlled. I will refer to the different origami used here as R09, R20 and R35 to indicate the nominal number of attached fluorophores per origami and the emission wavelength of the fluorophore used (R - red, labeling with ATTO647N). Before applying DNA origami as in vitro reference samples for benchmarking `quickPBSA`, I sought to confirm the labeling efficiency of 70 % stated by the manufacturer (GattaQuant, personal communication) with the emitter counting microscopy technique CoPS (see section 2.3) as well as intensity-based measurements.

I immobilized origami via biotin-streptavidin linkage as described above for tetraATTO647Nprobes. After deposition of origami on cleaned coverslips, individual origami appeared as bright, immobile spots in epifluorescence microscopy images (Fig. 4.8b). Before proceeding with CoPS measurements, I confirmed that the photostability of ATTO647N attached to DNA origami and imaged in PCD ROXS buffer was similar than previously observed for tetraATTO647Nprobes. For this, I acquired timelapse image series under illumination intensities comparable to intensities used for photostability measurements with tetraATTO647Nprobes (~2.4 $\text{kW}\,\text{cm}^{-2}$) and extracted the integrated frame-wise intensity around all PSFs visible in the first frame of an image series. I observed that R09 origami (nominally labeled with ATTO647N) exhibited a photostability comparable to that of tetraATTO647N probes in PCD ROXS buffer. R20 and R35 origami exhibited a reduced photostability with $t_{1/2}$ of ~50 % of that observed for R09 and tetraATTO647Nprobes under the same conditions (Fig. 4.8c,d). While the cause for this observation is unclear, one potential reason is a decreased interfluorophore distance for R20 and R35 (GattaQuant, personal communication). Since the photostability for all three types of origami was still 15-30-fold increased compared to measurements with tetraATTO647N in PBS, I performed all further measurements with DNA origami in PCD ROXS buffer.

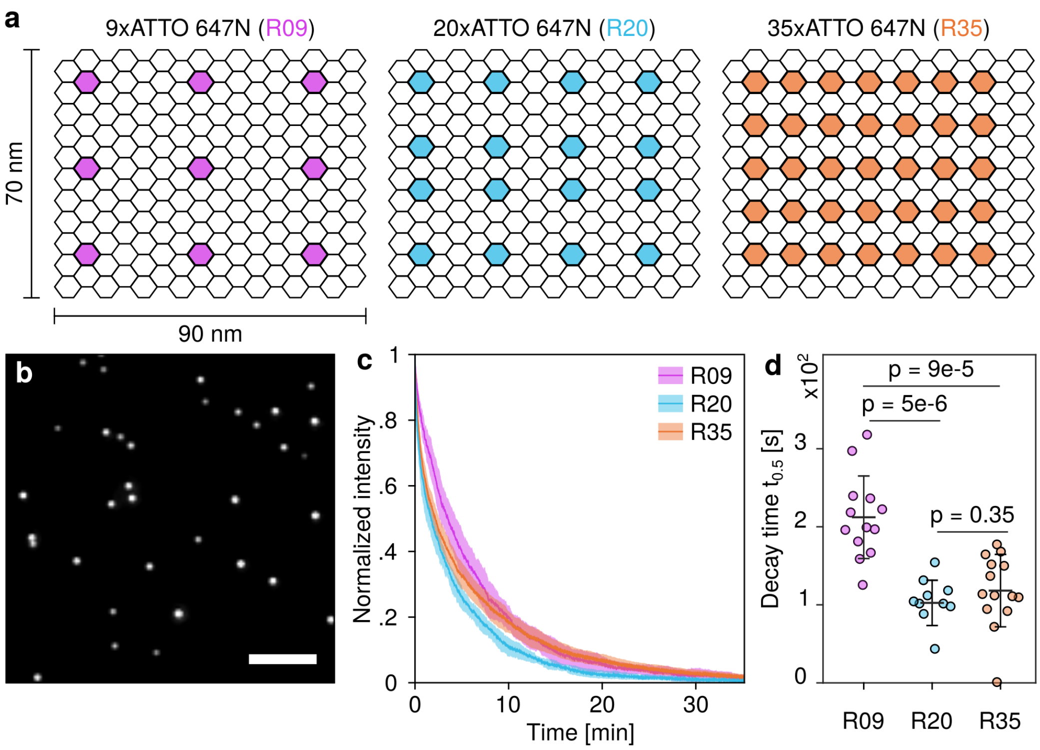

Figure 4.8 Design and photobleaching behavior of DNA origami labeled with varying number of fluorophores. R09, R20 or R35 origami were immobilized at single-molecule density via biotin-streptavidin linkage and images via widefield fluorescence microscopy in ROXS PCD buffer. **a,** Schematic layout of DNA origami templates with approximate physical size of of origami as indicated. Please note that the shown distribution of binding sites across origami was scrambled to protect intellectual property by the manufacturer. **b,** Representative image of a R35 origami sample. Scale bar: 5 µm. **c,** Averaged photobleaching curves for each origami type. Mean±SD from 10-14 FOVs per sample. **d,** $t_{1/2}$ obtained from bleaching curves of individual FOVs. Error bars: Mean±SD. Samples were compared using Welch's *t*-test.

CoPS has been used before to characterize the number of fluorophores attached to DNA origami[130] and a previous study established criteria for performing robust emitter counting measurements with CoPS[198]. I followed these previously established guidelines and performed measurements to determine the label number distribution per DNA origami for R20 origami. After removing individual measurements, which did not yield the required minimum molecular brightness (Fig. 4.9a), a unimodal label number distribution was obtained. This distribution was well described by a normal distribution with a mean of 14.6 ± 0.6 (95 % CI) corresponding to a mean labeling efficiency of 73 % (Fig. 4.9b). This is in excellent agreement with the expected mean fluorophore number per origami of 14 based on the provided labeling efficiency of 70 %. The standard deviation of the obtained label number distribution determined with CoPS was 5.0 ± 0.4, which is approximately 2.4-fold higher than expected at 70 % labeling efficiency (Fig. 4.9b, bars above distribution).

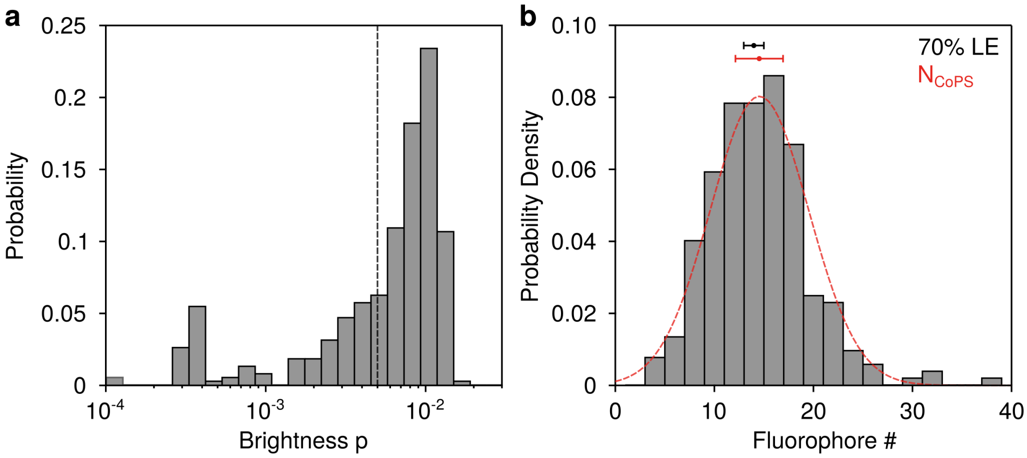

Figure 4.9 CoPS analysis of fluorophore numbers on R20 DNA origami samples. Biotinylated DNA origami were immobilized via biotin-streptavidin linkage and imaged in ROXS PCD buffer supplemented with 10 mM $MgCl_2$ 60 min after buffer addition. 386 pooled traces from 10 positions from 1 experiment. See methods (Section 6.2.2 for details). **a,** Distribution of brightness p obtained from CoPS fitting. Cut-off for trace filtering indicated by dashed line. **b,** Fluorophore number per origami after filtering for a minimum brightness of $\geq 5 \times 10^{-3}$. Distribution was modeled with a Gaussian function (dashed red line). Mean±SD from CoPS measurements (black) and from binomial distribution assuming 70 % LE (red) shown as dot and horizontal bars above distribution.

After validating the labeling efficiency of the R20 origami, I next investigated if DNA origami with different nominal numbers of fluorophores per origami exhibit the expected difference in single-molecule brightness. Assuming binding sites on origami are independent and labeling procedures across different batches of origami are comparable, it would be expected that the mean intensity of origami linearly scales with the number of binding sites per origami. To assess this, I imaged immobilized R09, R20 and R35 origami and localized signals from origami in images using the `thunderSTORM` emitter localization framework (Section 3.4.1). Amplitudes from sub-pixel localization of individual origami were then used to compare the different origami types based on their emission intensity. For all three origami types, I observed unimodal distributions, which could be well described with a Gaussian function (R^2>0.9 for all distributions) indicating that origami labeling had occurred evenly across origami (Fig. 4.10a). I then compared the mean intensity obtained for each origami type against the nominal number of fluorophores per origami and found that the mean intensity depended linearly on the nominal number of fluorophores per origami (R^2>0.99, Fig. 4.10b).

To compare the spread of recorded intensity distributions for each origami type with the expected spread based on the specified labeling efficiency and the nominal number of fluorophores for each origami type, I divided origami intensities by the slope of the linear fit shown in Fig. 4.10b to convert intensities into apparent fluorophore numbers. Based on this analysis and similar to the previous observations with CoPS, the spread of the intensity distribution for each origami type was between 1.9- (R09) and 4.0-fold (R35) higher as would be expected based on a labeling efficiency of 70 %.

Taken together, these results show that all three types of DNA origami evaluated in the scope of the presented experiments conform with the specified labeling efficiency provided by the manufacturer.

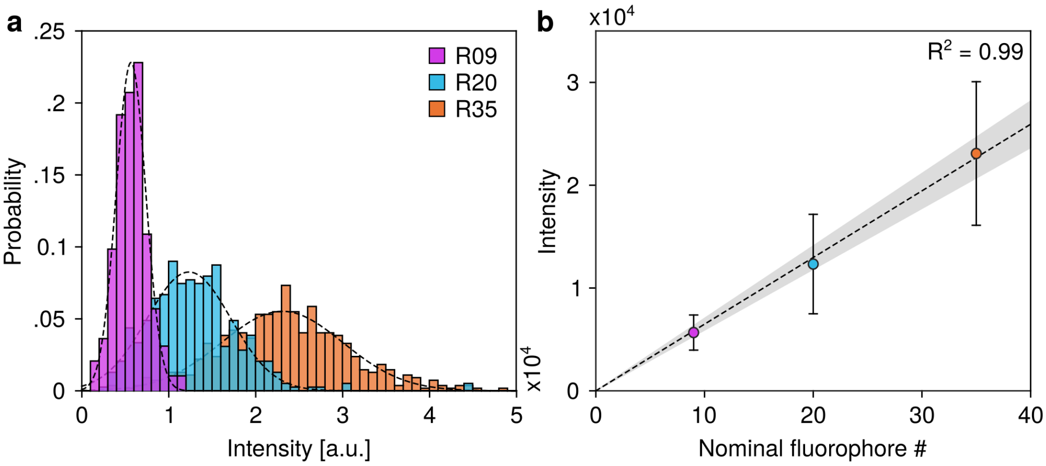

Figure 4.10 Characterization of DNA origami with 9,20 and 35 binding sites for ATTO647N-labeled oligonucleotides based on fluorescence intensity. DNA origami were immobilized via biotin-streptavidin linkage and imaged by epifluorescence microscopy. Individual origami were identified based on thunderSTORM peak localization and subsequent filtering of candidates based on their PSF width (100-200 nm). Intensities of individual origami were obtained from thunderSTORM localization and corrected for pixel size. **a,** Intensity distributions obtained for each origami type. Distributions were modeled with a Gaussian function (dashed lines). **b,** Mean intensity (circles) per origami against nominal number of fluorophores per origami. Error bars indicate standard deviation obtained from Gaussian modeling shown in a. Linear fit to mean intensity (dashed line) and corresponding fitting error (95 % confidence intervals, shaded region) show linear dependence of mean origami intensity and nominal number of fluorophores per origami. 198 to 546 individual spots per origami type were analyzed.

4.4.3 PBSA on DNA origami with up to 35 fluorophores

After having identified improved data acquisition conditions and after having verified the R09, R20 and R35 DNA origami as suited reference targets, I assessed the performance of `quickPBSA` in comparison with the previously described $\mathrm{KV}_{\mathrm{only}}$ and local intensity approaches for PBSA. For this, I acquired photobleaching datasets from all three origami types in ROXS PCD buffer and processed the obtained data with all three algorithms.

Similar to the intensity comparison of the same set of DNA origami described in the previous section, I found unimodal fluorophore number distributions for all three origami types when using `quickPBSA`. At the same time, the widths of the emitter number distributions increased with increasing nominal fluorophore numbers (Fig. 4.11a). I further observed a linear dependence of mean emitter numbers measured with `quickPBSA` on the nominal fluorophore number. The linear fit across all three mean emitter numbers revealed that `quickPBSA` exhibited a slight tendency (<1 emitter) towards underestimating the nominal fluorophore number (Fig. 4.11b). The local intensity approach yielded similar results confirming the linear increase in emitter numbers with the increasing number of nominal fluorophore numbers (Fig. 4.11c). In comparison to `quickPBSA`, and the local intensity approach, the $\mathrm{KV}_{\mathrm{only}}$ step detection algorithm showed a systematic trend towards underestimating emitter numbers with increasing nominal fluorophore numbers.

All three PBSA approaches tended to overestimate the emitter number variability compared to the estimated variability at 70 % labeling efficiency. While for R09 origami the majority of variability in the emitter number distribution could be explained with the expected variation due to imperfect labeling, the recovered standard deviations for the origami with higher label numbers (R20 and R35) were substantially larger than expected (Fig. 4.11d). Here, the variability in emitter number estimates was highly similar for `quickPBSA` and the local intensity approach. Both approaches achievd a slightly higher emitter number variability for R20 in comparison to the variability achieved with CoPS (Fig. 4.9), but achieved narrower distributions for emitter numbers when based on intensities alone (Fig. 4.10). I previously observed that for semi-synthetic data, filtering of traces after preliminary step detection resulted in $\sim$17 % of traces being removed (Fig. 4.5). In the case of data acquired on DNA origami, only a weak decrease in the fraction of successfully processed traces to $\sim$70 % for R35 origami was observed.

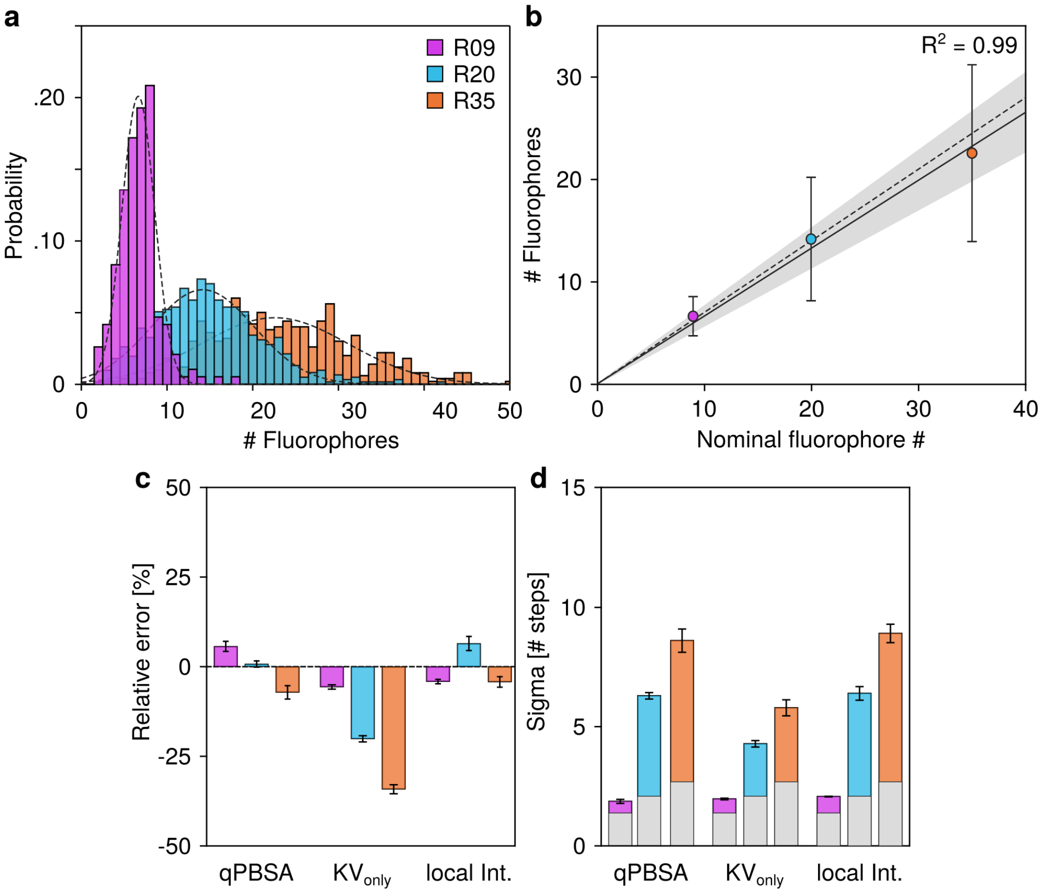

Figure 4.11 quickPBSA robustly reports the number of fluorophores per DNA origami. DNA origami with 9, 20 or 35 binding sites for ATTO647N-labeled imager strands were immobilized on coverslips via biotin-streptavidin linkage and imaged in ROXS PCD buffer. Positions of origami were determined using thunderSTORM and traces were extracted and processed using the `quickPBSA` framework. **a,** `quickPBSA` results for each DNA origami class as indicated. Distributions for each DNA origami class were modeled with a Gaussian function (dashed lines). See Table B.3 for information about sample size and results from Gaussian modeling. **b,** The observed mean fluorophore number per DNA origami class depends linearly on the nominal number of fluorophores per origami. Results from all three DNA origami classes were modeled with a linear fit (black line, fitting uncertainty - shaded region). Expected mean numbers assuming 70 % labeling efficiency indicated by dashed line. Errorbars indicate sigma from respective Gaussian fit shown in a. **c,** Bias of fluorophore number estimates for different algorithms. Means from Gaussian fits normalized by nominal number of fluorophores assuming 70 % labeling efficiency. Errorbars: 95 % CI from Gaussian fits. **d,** Precision of fluorophore number estimates for different algorithms. Sigma from Gaussian fits. Expected variance due to limited labeling efficiency indicated in gray. Errorbars: 95 % CI from Gaussian fits.

Based on the observation made above, I concluded that `quickPBSA` is suited for emitter counting in a simplified in vitro scenario where all targets were located on the surface of coverslips and were immobilized at low densities. In line with previous observations from the analysis of semi-synthetic data, the local intensity approach yielded comparable results to `quickPBSA`, while the KV_{only} approach performed considerably worse on origami with nominal emitter numbers $\geq$9. I next sought to demonstrate that `quickPBSA` performs well on in vitro samples in general by testing if data acquired on different microscopes and with different fluorophores attached to DNA origami resulted in a similar performance.

4.4.4 Improved throughput via data acquisition with large field of view

In the previous section, I showed that `quickPBSA` can be used for emitter number estimation from experimental data acquired on in vitro samples using a setup with a Gaussian illumination profile and a Electron-multiplying CCD (EMCCD) camera for image acquisition (setup A). For further validation of `quickPBSA`, I acquired photobleaching data from R09 DNA origami on a second microscope (setup B) equipped with a sCMOS camera for image acquisition. Setup B employed a refractive beam shaping device to homogenize the excitation intensity across the illuminated area. To study the effect of flatfielded illumination on emitter intensities, I imaged DNA origami samples randomly at >60 positions and localized individual diffraction-limited spots via SMLM analysis. The amplitudes of obtained localizations were used to determine the relative excitation power density in dependence of the origami position within the FOV for each setup (Fig. 4.12a). The normalized sector-wise intensities for setup A revealed a strong positional dependence of detected signal from individual origami with differences >50 % across the FOV (Fig. 4.12b). In contrast, the signal variations from individual origami on setup B showed substantially lower variability and no systematic trends across the FOV. This suggests that `quickPBSA` might perform better on data acquired from setup B since this setup introduces less trace-to-trace variation in single step heights during PBSA than setup A. In addition to this putative advantage in the quality of raw traces, the sCMOS detector used in setup B also enabled image acquisition with a ~16.5-fold larger FOV.

This resulted in a largely (>11x) increased number of photobleaching traces that could be recorded from an identical sample on setup B, where a typical dataset contained $\sim 5 \times 10^2$ individual traces/FOV. I compared the results obtained by `quickPBSA` on raw data from setup A and B by acquiring a dataset comprised of $>2 \times 10^3$ traces from FOVs and processing extracted traces with `quickPBSA` as described above. The estimated mean emitter number per R09 origami of 6.0 ± 2.2 (mean±) obtained from data acquired on setup B exhibited a comparable deviation from the expected emitter number of 6.3 based on the previously determined labeling efficiency as for the data acquired previously on setup A. Interestingly, also the fraction of successfully processed traces of 75 % was similar as in the previously processed dataset from setup A. This suggests, that the technical improvements included in setup B helped to increase the number of photobleaching traces acquired in a given time, but did not result in narrower emitter number distributions after PBSA analysis. `quickPBSA` therefore seems to perform robustly on data acquired with different optical setups and, as expected, somewhat independent of raw data uniformity.

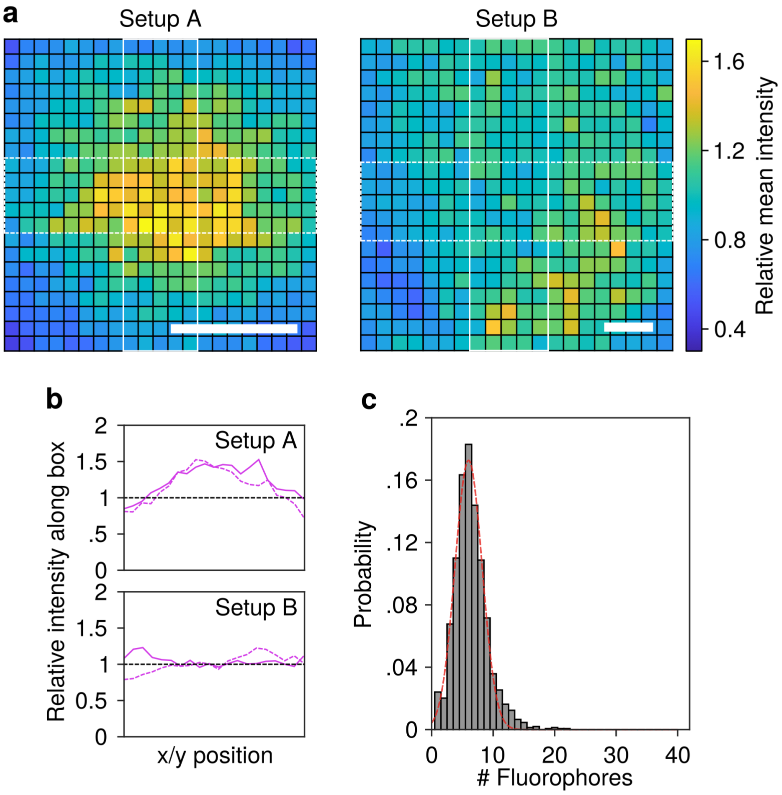

Figure 4.12 Characterization of setups A and B and quickPBSA analysis using data acquired on setup B. Illumination profiles were characterized using images acquired on R20 origami samples, sector-wise mean intensities of origami were obtained by SMLM analysis. On average, 165 ± 23 and 1 410 ± 67 (mean±SD) localizations were detected per sector on setup A and setup B, respectively. The relative sector-wise mean intensity was computed by dividing sector-wise mean intensities by the median intensity across all sectors. **a,** Normalized mean intensity within each sector across FOV. Scale bars: 20 µm. **b,** Intensity profiles along horizontal (dashed lines) and vertical (solid lines) regions indicated in a. For each profile column- (horizontal) or row-wise (vertical) means across sectors contained in boxes were computed. **c,** Distribution of fluorophore numbers determined using `quickPBSA` on photobleaching traces recorded using setup B. Distribution of fluorophore number estimates was modeled using a Gaussian function (dashed red line). Data was acquired jointly with Jonas Euchner.

4.4.5 Importance of fluorophore choice

In addition to ensuring robustness across different microscopes used for data acquisition, it was also important to test how well `quickPBSA` could cope with data acquired on samples labeled with different fluorophores. I therefore decided to evaluate how well emitter counting with `quickPBSA` worked on DNA origami samples labeled with the fluorophore ATTO565, which exhibits a molecular brightness comparable to that of ATTO647N (Table B.4). To distribute the available photon budget for ATTO565 to a roughly comparable number of data points (frames) as in previous measurements with origami labeled with ATTO647N, I adjusted the exposure time per frame and the illumination intensity in a way which resulted in similar photobleaching rates per frame (Fig. 4.13,a,b). PBSA with `quickPBSA` on data acquired on Y09 origami resulted in an unimodal emitter number distribution, but showed a pronounced tail toward large emitter number (Fig. 4.13c). The obtained emitter number per Y09 origami of 7.5 ± 1.5 (mean±SD) was ∼20 % higher than expected for a labeling efficiency of 70 %, while the fraction of analyzable traces was similar (∼80 %) as for origami labeled with ATTO647N.

Closer inspection of photobleaching traces extracted from timelapse data revealed that both fluorophores exhibited more than one fluorescent state with different brightness. This behavior was distinguishable from transient dark states, i.e. binary switching between a fluorescent and a non-fluorescent state, since switching of individual fluorophores between both states was observed before single-step bleaching to background levels (Fig. 4.13d,e). While the observed changes in brightness were rather frequent for ATTO647N and exhibited a lower intensity difference ΔI of 24 %, brightness changes occurred less often for ATTO565 and exhibited a substantially larger ΔI of 63 %. If the final photobleaching step in a trace occurs from the lower brightness state in a significant fraction of analyzed trace, the mean last step height will be underestimated which in turn would lead to an increase in spurious steps detected in traces. This observed behavior and the subsequent influence on the emitter number estimates obtained through PBSA with `quickPBSA` highlights the importance of label selection and validation of the experimental system using reference standards with known stoichiometry.

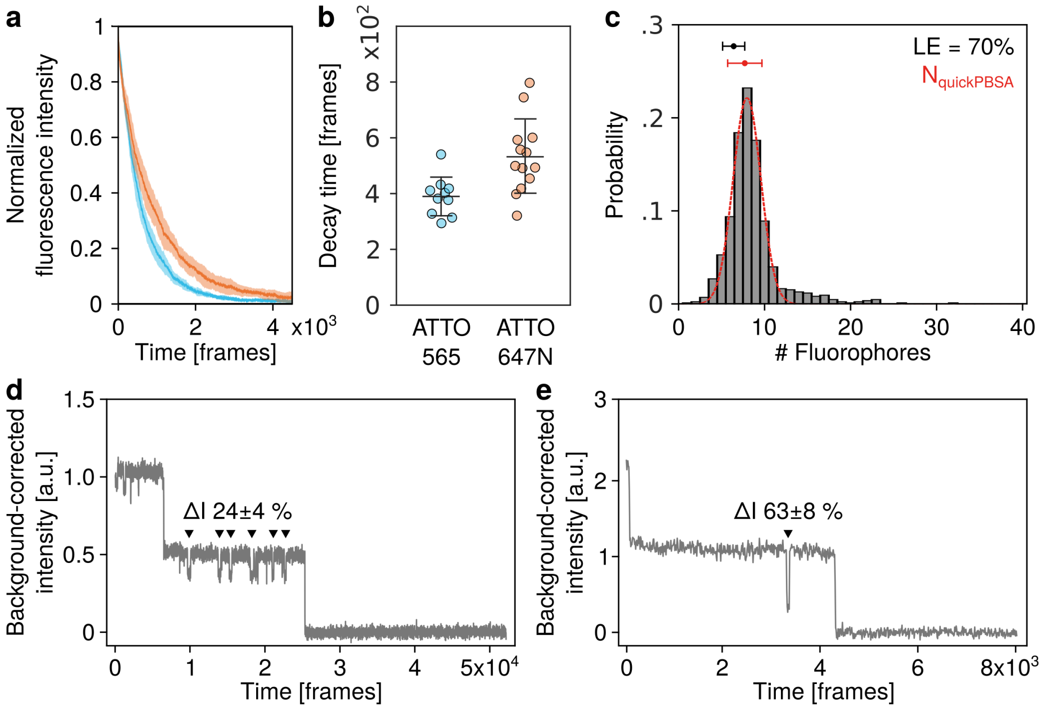

Figure 4.13 Characterization of ATTO565 for PBSA with quickPBSA. DNA origami with 9 binding sites for ATTO565-labeled imager strands (Y09) were immobilized on chambered coverglasses and imaged in ROXS PCD buffer. **a,** Normalized photobleaching curves for DNA origami labeled with ATTO565 (blue) or ATTO647N (orange). Mean (solid line) and SD (shaded region) across ≥12 FOVs per fluorophore. **b,** Corresponding $t_{0.5}$ values for each FOV. Error bars: mean±SD across FOVs. Samples were compared using Welch's *t*-test. **c,** Fluorophore number estimate distribution obtained from quickPBSA analysis of photobleaching traces from Y09 samples. The obtained distribution was modeled with a Gaussian function (dashed red line). Expected (black) and measured (red) fluorophore number distribution for the sample assuming LE=70 % indicated by error bars above histogram. See Table B.3 for full summary statistics of dataset. **d,e,** Representative intensity traces recorded from DNA origami labeled with ATTO647N (d) and ATTO565 (e) imager strands exhibiting multiple bright states followed by a single final bleaching step. The intensity difference between both bright states, ΔI (mean±SD), was estimated by manually annotating traces (n=6) for each fluorophore.

4.5 In situ emitter number measurements with quickPBSA

In the last two sections, I described the design principles of `quickPBSA` and showed that `quickPBSA` is suited for emitter number estimation on semi-synthetic data. I also showed that after optimization of data acquisition conditions, `quickPBSA` achieved results in excellent agreement with alternative emitter counting microscopy techniques for data acquired in vitro on DNA origami reference standards. Based on these results, an accessible counting range for `quickPBSA` of up to $\sim$35 emitters contained in a single structure could be verified experimentally.

However, several experimental parameters in the vitro experiments described above differ from conditions typically encountered in cell biological samples. For example, the DNA origami used as target were deposited on clean coverslips. This resulted in samples where all clusters where located in the same plane and consequently exhibited relatively low variations in signals between individual origami. Furthermore, DNA origami samples were prepared at a low origami density which made identifying individual origami and extracting intensity traces a straightforward task. This also resulted in negligible crosstalk between the signals from individual origami. In contrast, cellular samples typically exhibit a complex three-dimensional organization with protein complexes being located in different planes and at varying distances to each other. In addition, signal contributions from complexes located in out-of-focus planes and from complexes located in close proximity to each other will have to be accounted for. This complicates the localization of individual complexes and makes background correction an essential step during trace extraction. Finally, the fluorophore used for labeling DNA origami, ATTO647N, is unsuited for labeling of targets in cellular samples due to its lipophilicity, which causes fluorescent labels conjugated with ATTO647N to bind non-specifically to cellular membranes[205].

In this section, I will describe how I addressed these challenges to establish `quickPBSA` as a tool for counting of POI copy numbers in protein complexes within cells. In the first part of the section, I will address the need for alternative labels for in situ protein counting by PBSA, by comparing different labels with respect to their brightness, photostability and tendency to populate dark states. In the second part, I will describe how I established and tested two candidate systems which were recently reported as in-cell reference standards for benchmarking and calibration of emitter counting microscopy techniques. In the third part of this chapter, I will describe how the `quickPBSA` framework was extended to meet the demands posed by the complexity encountered in cellular samples. Finally, I will show that after adaption of the data acquisition and data analysis workflow, `quickPBSA` was successfully used to measure protein copy numbers in cellular complexes within cells.

4.5.1 Optimized data acquisition for in situ counting

Ideally, protein counting by fluorescence microscopy should be indifferent to the way a POI is fluorescently labeled. This way, the easiest possible way to introduce a fluorescent label, with minimal perturbations of the biological system under study could be chosen. In practice, however, the properties of fluorescent labels play important roles when selecting a fluorescent label (section 2.2). For emitter counting by PBSA, the brightness, photostability and the tendency to exhibit photoblinking are critical to obtain photobleaching traces with high SNR, well-separated photobleaching steps and few signal fluctuations due to photoblinking. To identify a fluorescent labeling strategy well-suited for PBSA and compatible for labeling in cells, I compared different fluorescent labels with respect to their photostability and their tendency to exhibit photoblinking. The set of compared fluorophores included the three fluorescent proteins EGFP, mCherry and mNeonGreen and the self-labeling protein tags $\mathrm{SNAP_f}$-tag and HaloTag, which each were evaluated with their respective -SiR and -TMR ligands. EGFP and mCherry were included since both have been used extensively in previous PBSA experiments[117]. mNeonGreen was recently reported as fluorescent protein with improved brightness[206] and could therefore confer advantages for PBSA. $\mathrm{SNAP_f}$-tag and HaloTag were included since both tags bind small organic fluorophores, which tend to exhibit better photostability than fluorescent proteins and are accessible to further enhancements with photostabilizing buffers.

The brightness of fluorescent labels has been extensively studied in the literature and spectroscopic characteristics of fluorophores are often available from databases[64] or in primary literature[183]. I therefore compared the brightness of labels based on molar extinction coefficients (ε), quantum yields (Φ) reported in the literature. Whenever possible, I used values measured in aqueous buffer and, in the case of protein tag ligands after conjugation to the respective protein tag. Table B.4 provides a detailed overview of values used for calculation and respective references to published work. To correct for the influence of excitation light source on the molecular brightness, I determined a correction factor $CF = \lambda_{ex}/\lambda_{abs}$, which corrects the nominal brightness for mismatches between the peak absorption wavelength λ_{abs} and the excitation wavelength λ_{ex}. The effective molecular brightness b' is then computed as $b' = CF\varepsilon\Phi$.

As shown in Fig. 4.14, the nominal brightness of the different fluorophores tested in this work varied by more than six-fold between mCherry and BG-SiR, which exhibited the lowest molecular brightness and mNeonGreen (mNG) which exhibited the highest molecular brightness among the tested fluorophores for in-situ PBSA. The nominal molecular brightness of mNG approached the brightness of ATTO647N previously used for in-vitro PBSA measurements, but limited spectral overlap between its excitation wavelength and the excitation light sources available in this work led to a reduced effective brightness. Among the ligands for self-labeling protein tags, TMR ligands for both tags exhibit a very similar effective molecular brightness after coupling to their respective tags, while for Silicon rhodamine (SiR), a strong influence of the different protein tags was reported recently[207].

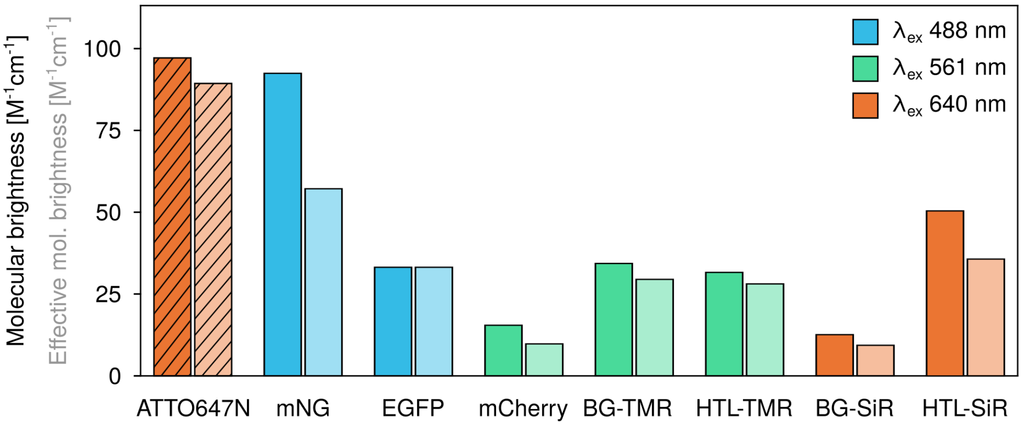

Figure 4.14 Nominal and effective molecular brightness for bio-compatible labels tested in this work. Molar extinction coefficients and QY were obtained from previous studies. The effective molecular brightness (light colors) at corresponding excitation wavelengths was computed from the nominal molecular brightness (dark colors) and a measured correction factor to account for spectral mismatches. Dashed fluorophores were not tested in cellular samples and are included for comparison to fluorophores used for in-vitro experiments. See Table B.4 for full list of fluorophore properties and corresponding references.

Next, I set out to determine the photostability of the panel of selected labels. For this, I followed the approach described above for ATTO647N (Section 4.4.1) to determine the time to 50 % intensity decrease, $t_{1/2}$. To mimic the conditions typically encountered in a PBSA experiment in cells, I expressed each label as fusion construct with a native protein to target the label to a known subcellular structure and acquired data after chemical fixation of cells (Fig. 4.15a). Here, fusion of labels to proteins with known subcellular localization served as validation control to identify cells successfully transfected and to ensure none of the labels resulted in aberrant subcellular localization (Fig. 4.15b-e). Because fixed and permeabilized cells are permeable to buffers, I assumed that despite the fact that fluorescent proteins and protein tags were conjugated to different subcellular targets, the environment in which the fluorophores reside can be assumed to be comparable.

Each fluorophore was characterized in each of the four buffers also used for in vitro experiments with ATTO647N, which were: 1xPBS, ROXS $NaSO_3$, ROXS PCD, ROXS God-Cat. To correct for differences in excitation power density used for excitation of fluorophores with different wavelengths, I corrected the measured $t_{1/2}$ to $1\,\mathrm{kW\,cm^{-2}}$ assuming a linear relationship between bleaching rate and excitation power density.

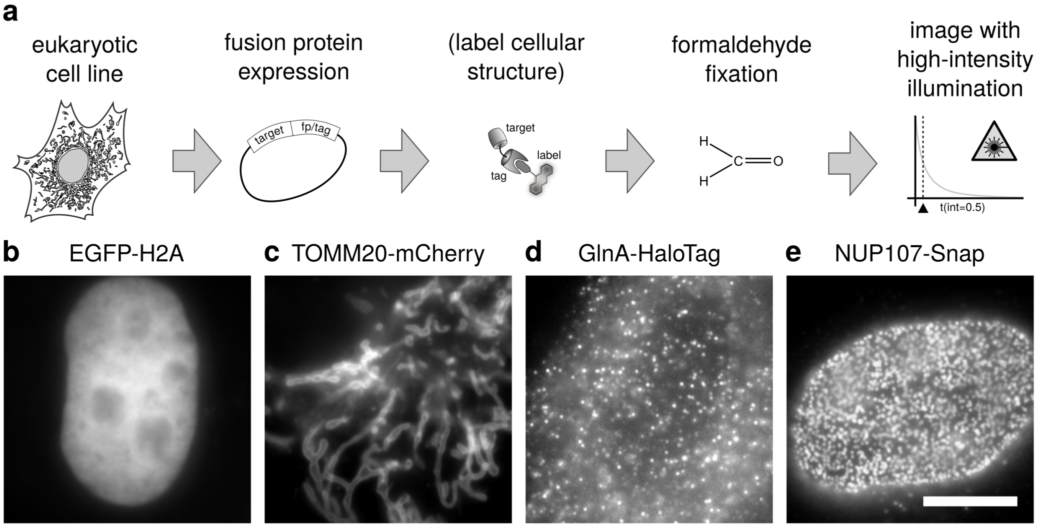

Figure 4.15 Assay for in situ photostability characterization. a, Eukaryotic cell lines transiently or stably expressing proteins fused to either fluorescent proteins or protein tags were used to determine fluorophore photostability in situ. Samples expressing protein tag fusion proteins were labeled with fluorescent substrates for protein tags prior to PFA fixation. Samples expressing fusion proteins labeled with fluorescent proteins were fixed with PFA without additional washing or incubation steps. Samples were then imaged in PBS or ROXS buffers under high intensity illumination matching the absorption spectrum of the corresponding fluorophore. **b-e,** Representative images from cells expressing indicated fluorescent protein or protein tag fusion proteins. For TOMM20-mNeonGreen a comparable subcellular localization as for TOMM20-mCherry was observed. Scale bar: 10 µm.

Photobleaching under high intensity illumination ($0.4\,\mathrm{kW\,cm^{-2}}$), resulted in a fast loss of signal intensity for fluorescent proteins within seconds and exchanging PBS for photostabilizing buffers resulted in a further decrease in photostability for EGFP and mCherry (see e.g. Fig. 4.16a for behavior of EGFP). In PBS, EGFP exhibited the highest photostability $t_{1/2}$: 2.2 ± 0.2 s), with mNeonGreen and mCherry exhibiting $t_{1/2}$ between 0.6 ± 0.2 and 1.2 ± 0.1 s (Fig. 4.16b). While EGFP showed comparable stability in all three ROXS buffers, the negative effect on photostability for mCherry was observed to be stronger in ROXS buffer with enzymatic oxygen removal, possibly due to a stronger sensitivity to changes in pH which have been reported for both, the PCD and GodCat system[28]. Interestingly, I observed no decrease in photostability for mNeonGreen in ROXS buffers. This could be due to improved stability of mNeonGreen protein structure[206]. Nevertheless, mNeonGreen was found to bleach about 3-4-fold faster than EGFP in PBS. As previously observed for ATTO647N, all three fluorescent proteins exhibited bi-exponential intensity decays in all tested buffers again indicating the complex nature of photobleaching (see Fig. B.2).

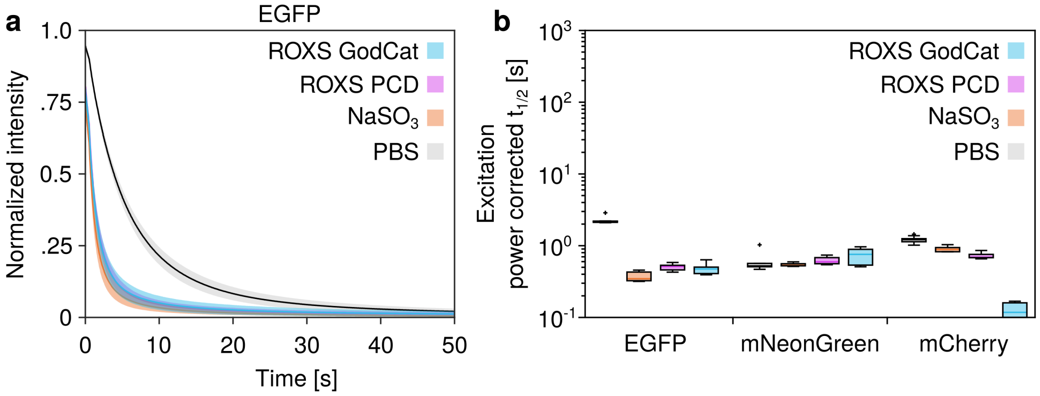

Figure 4.16 Photostability of selected fluorescent proteins in ROXS buffers. HeLa or COS-7 cells expressing fusion proteins with EGFP, mNeonGreen or mCherry (see 4.15) were seeded into chambered coverglasses, fixed with 3.7 % PFA and imaged by widefield fluorescence microscopy in indicated buffers. 488 nm (EGFP, mNeonGreen) or 561 nm (mCherry) excitation light at 0.4 kW cm^{-2} and 0.52 kW cm^{-2} average power density respectively were used for imaging. The intensity over time was recorded from ROIs obtained by applying an intensity threshold (Section 6.3.6). For each condition, 7-10 FOVs from two independent experiments were recorded. **a,** Intensity decays from each FOV recorded from EGFP samples were normalized to the intensity at time zero and averaged over all FOVs (lines). The standard deviation across FOVs for each condition is shown as shaded region. **b,** Time to half maximum intensity ($t_{0.5}$) corrected for excitation power for EGFP, mNeonGreen and mCherry across all ROXS buffer systems tested.

In line with the intensity decay pattern observed above for fluorescent proteins, Tetramethyl rhodamine (TMR) and SiR substrates for SNAP$_f$-tag and HaloTag also exhibited bi-exponential signal decays upon high intensity illumination in all tested conditions (Fig. B.3). Extracting $t_{1/2}$ from the signal decay of TMR and SiR in PBS showed that the photostability of TMR was comparable to EGFP, while the photostability of SiR was 2-3-fold higher. In contrast to what I observed for fluorescent proteins, all three tested ROXS buffers had a pronounced stabilizing effect on both, TMR and SiR. For example, BG-SiR showed 2-fold increased photostability in ROXS NaSO$_3$ and 10- and 20-fold increased photostability in ROXS GodCat and ROXS PCD, respectively (Fig. 4.17a). Both dyes exhibited similar photostabilities after conjugation to SNAP$_f$-tag and HaloTag in PBS and in ROXS buffers.

The effect of ROXS buffer and enzymatic oxygen removal was more pronounced for SiR than for TMR (Fig. 4.17b). Overall, SNAP$_f$-tag conjugated with BG-SiR in ROXS PCD buffer exhibited the highest photostability with an excitation power corrected $t_{1/2}$ of 120 ± 36 s which represents a 54-fold improvement over EGFP in PBS.

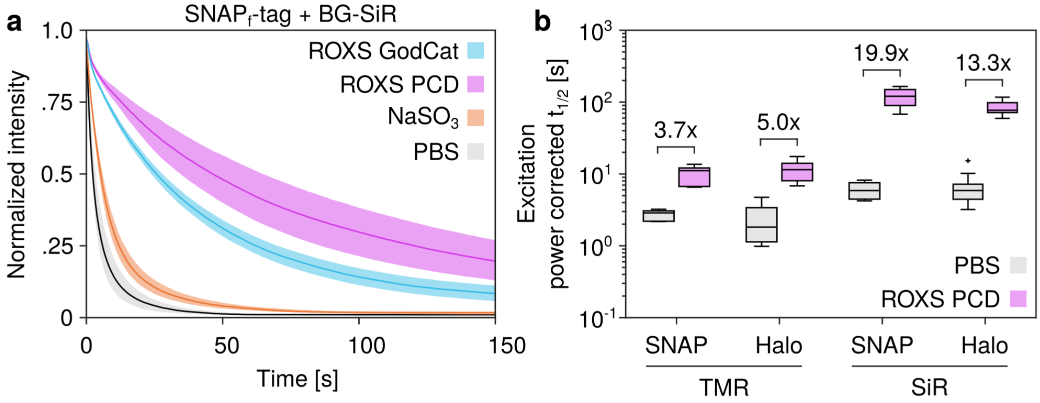

Figure 4.17 Photostability of TMR and SiR conjugated to SNAP$_f$-tag or HaloTag in ROXS buffers. HeLa or U2OS cells stably expressing fusion proteins with SNAP-tag or HaloTag (see 4.15) were seeded into chambered coverglasses stained with BG-TMR/SiR (100 nm) or HTL-TMR/SiR (200 nm) for 120 min, washed and fixed with 3.7 % PFA. Samples were then imaged in indicated buffers by widefield microscopy at 0.52 kW cm^{-2} (561 nm, BG/HTL-TMR) or 1.26 kW cm^{-2} (640 nm, BG/HTL-SiR) average excitation power density. The intensity over time was recorded from ROIs obtained by applying an intensity threshold (Section 6.3.6). For each condition, 7-10 FOVs from two independent experiments were recorded. **a,** Exemplary intensity decays from each FOV recorded from SNAP-tag samples labeled with BG-SiR were normalized to the intensity at time zero and averaged over all FOVs (lines). The standard deviation across FOVs for each condition is shown as shaded region. **b,** Effect of ROXS PCD on photostability. Excitation power-corrected time to half maximum intensity ($t_{0.5}$) for PBS and ROXS PCD and each fluorophore/tag combination. Fold improvement in photostability (median vs. median) as indicated.

For a global comparison of the tested labels and imaging buffers, I compared the reported molecular brightness against the measured photostabilities (Fig. 4.18). From this graph, it is clear that none of the labels suited for in-cell labeling via genetic fusion constructs achieves the combined molecular brightness and photostability of ATTO647N. It is also clear that the variability in photostability 1×10^3 across the tested fluorophore - buffer combinations is substantially larger than the differences in molecular brightness (6-fold). Since the excitation power densities used for the measurements presented above are at least one order of magnitude below fluorescence saturation power densities, brightness and correspondingly, SNR, can be traded against reduced photostability by increasing excitation power. I therefore reasoned that using more photostable rather than brighter fluorophores represents a promising strategy for optimizing data acquisition for PBSA. Following this reasoning, Fig. 4.18 suggests that fluorophore conjugates for SNAP$_f$-tag and HaloTag labeling might be better suited as labels compared to fluorescent proteins. Among the tested fluorophore ligands, SiR conjugates showed a 10-100-fold higher photostability compared to TMR conjugates and would therefore be preferred. The overall low photostability of fluorescent proteins also indicates that for PBSA on complexes containing large fluorophore numbers, interpretation of photobleaching trace will be challenging when using fluorescent proteins as labels.

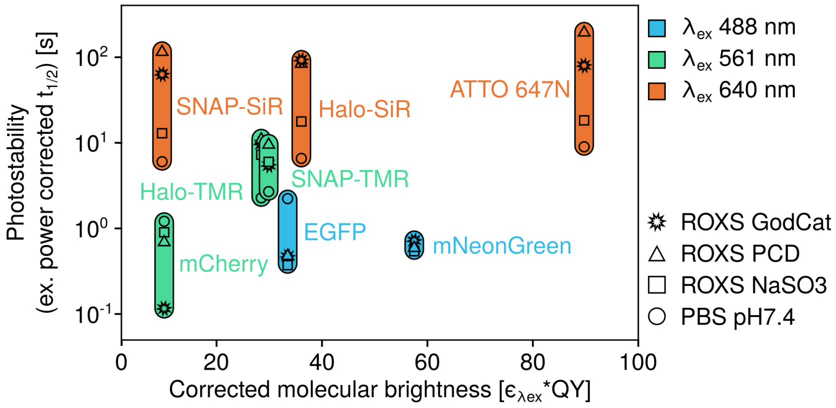

Figure 4.18 Overview of fluorophore brightness and mean photostability for all tested fluorophore and buffer combinations. Fluorophore brightness at indicated excitation wavelengths. Photostability was measured at different excitation power densities depending on excitation wavelength and normalized to $1\,\mathrm{kW\,cm^{-2}}$ (see Section 6.2.5). Symbols and color coding according to applied buffer composition excitation wavelength.

In addition to photostability and molecular brightness, the propensity of fluorophores to undergo transitions into temporary dark states with millisecond to >second lifetime can strongly influence the interpretability of recorded photobleaching traces. To identify labels and buffers which allowed me to minimize these undesired dark state transitions, I performed a quantitative evaluation of photoblinking behavior. Since fluorescent proteins had already exhibited a very low photostability in the experiments described above, I only included SNAP$_f$-tag and HaloTag each with TMR and SiR ligands and EGFP. Further, as ROXS PCD had already been identified as the strongest photostabilizing buffer, I performed these experiments only using PBS (as control) and ROXS PCD buffers. For this set of 10 conditions, I measured the kinetic rate parameters for transitions between the fluorescent and dark states using a previously described approach based on image correlation[208]. As samples, I used the HeLa gSEP cell line described in chapter 3, which enabled me to obtain samples where SNAP$_f$-tag, HaloTag or EGFP were expressed at single-molecule density to ensure that signals emitted by EGFP or HTL/BG-fluorophore could be easily observed by TIRF microscopy after chemical fixation. On these samples, I acquired image series of 1×10^3 frames (Fig. 4.19a), computed the temporal Autocorrelation function (ACF) (Fig. 4.19b) and fitted the ACF with a simplified three-state model of the photoblinking and photobleaching process (Fig. 4.19c) using MATLAB code provided by Sehayek et al.[208]. This fit then provided the on- and off-rates, k_{on} and k_{off} for transitions between the fluorescent on and the non-fluorescent off state, which were used as basis to compare photoblinking behavior between conditions.

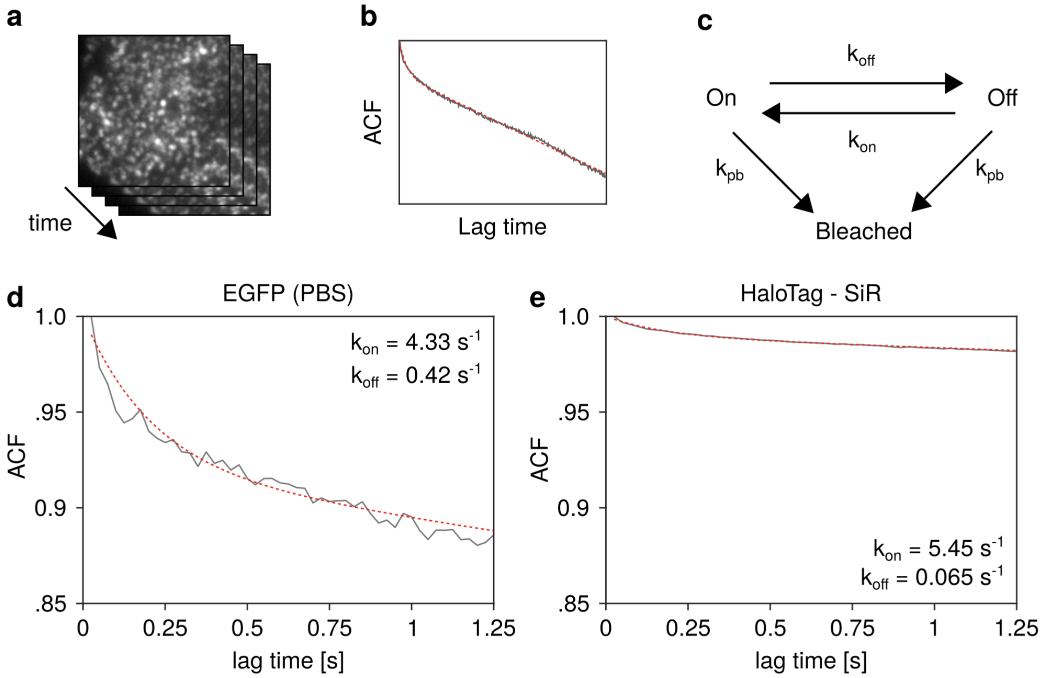

Figure 4.19 Principle of image correlation-based photoblinking characterization. a, Image series recorded on HeLa gSEP cells (a) were used as input data for computing the ACF (b). The obtained ACF was then fitted to a three state model of photoblinking and -bleaching with rates k_{off}, k_{on} and k_{pb} as fit parameters. **d,e** Representative fits (dashed) and raw ACFs firstplural (gray) for indicated combination of fluorophore and buffer exhibiting strong photoblinking (d) and weak photoblinking (e). Fitted rates are given for each fit. Data shown in (d,e) produced jointly with Jonas Euchner, Yin Xin Ho and Stanimir Asenov Tashev.

The chosen sample preparation and imaging appraoch yielded well-reproducible ACFs firstplural across replicates for the different combinations of label and buffer (Fig. B.4). Only for EGFP, a larger variability of ACFs firstplural computed from samples in PBS was observed. For EGFP in ROXS PCD buffer, no ACFs firstplural could be computed due to rapid photobleaching. ACFs firstplural computed from both, samples which exhibited strong and samples which exhibited weak photoblinking were well-described by the chosen fluorophore model (Fig. 4.19d,e).

In PBS, the mean on-time ($t_{on} = 1/k_{off}$) as well as the mean equilibrium constant ($K = k_{on}/k_{off}$) were similar for EGFP and SNAP$_f$-tag conjugated with BG-TMR or BG-SiR (Fig. 4.20). Exchanging the buffer for fluorophores conjugated to SNAP$_f$-tag resulted in an increase in t_{on} and K. For TMR- and SiR-ligands conjugated to HaloTag, higher mean k_{on} and K values were observed in both buffers. Interestingly, a decrease in k_{on} was observed upon switching from PBS to ROXS PCD buffer for both, HTL-TMR and HTL-SiR conjugated to HaloTag.

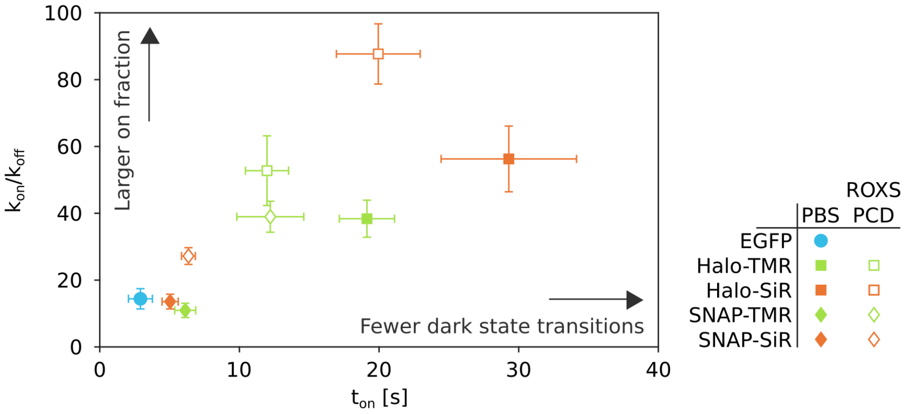

Figure 4.20 Overview of photoblinking tendency. Comparison between equilibrium constant K (k_{on}/k_{off}, mean±SD) and on-time t_{on} (mean±SD) obtained from image correlation experiments for the indicated conditions. Color coding indicates excitation wavelengths: 488 nm (blue), 561 nm (green) and 640 nm (orange). See Fig. B.5 for comparison of k_{off} and k_{on} for all tested conditions. 22-54 measurements from 2-4 independent experiments per condition. Data was produced jointly with Jonas Euchner, Yin Xin Ho and Stanimir Asenov Tashev.

The same qualitative behavior was also observed upon visual inspection of single-molecule intensity traces extracted from samples used for autocorrelation-based analysis of photoblinking using the `quickPBSA` candidate ROI detection and trace extraction modules. Intensity traces from individual EGFP molecules in PBS showed lower signal, frequent transitions between bright and dark states and tended to irreversibly photobleach within few seconds (Fig. 4.20a). In contrast, SiR conjugated to SNAP$_f$-tag (Fig. 4.20b) or HaloTag (Fig. 4.20c) exhibited a photostability similar to that of ATTO647N and only very few dark state transitions when imaged in ROXS PCD buffer (Fig. 4.20c).

To summarize, I compared different fluorescent labels suited for in-cell labeling which revealed that self-labeling protein tags in combination with SiR-conjugated ligands represent promising labels for in situ PBSA with `quickPBSA` and presumably also with other PBSA algorithms. While the brightness of SiR ligands was found to be 2.5- or 9-fold lower than ATTO647N, the photostability of both ligands was limited to 40-60 % of the stability of ATTO647N. In comparison, the photostability of EGFP was ~50-fold lower than that of -SiR ligands and EGFP exhibited a significantly increased propensity to undergo dark-state transitions, which makes them less suitable for in situ PBSA.

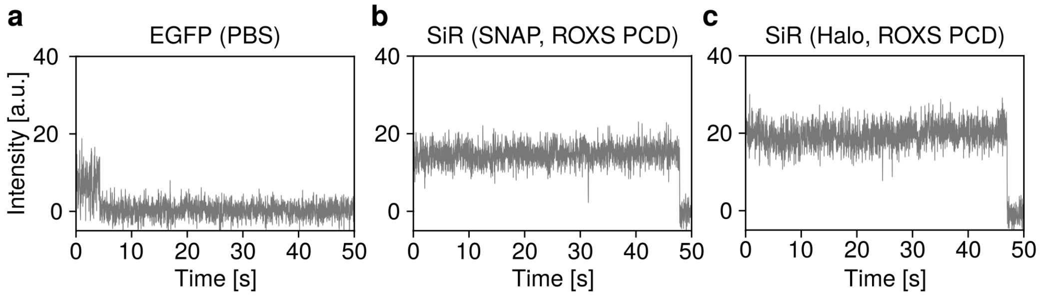

Figure 4.21 Representative single-molecule intensity traces. Traces for indicated fluorophores imaged in indicated buffers obtained by TIRF microscopy from chemically-fixed HeLa gSEP without (a) or after labeling with HTL of BG-SiR (b,c).

4.5.2 In cell reference target for benchmarking of PBSA approaches

In the previous part I showed that SNAP_f-tag or HaloTag in combination with photostable fluorescent ligands such as SiR and in particular when imaged using photostabilizing ROXS buffers allow for high-quality in situ data well-suited for PBSA. Importantly, intensity traces acquired under those optimized conditions in situ approach the quality of traces generated with ATTO647N in vitro. For the validation of `quickPBSA` in vitro described in section 4.4, I used DNA origami with varying numbers of binding sites as reference structure. However, DNA origami are unsuitable for in cell experiments as they require in vitro assembly from DNA oligonucleotides and labeling of DNA oligonucleotides with chemistry that is incompatible with site-specific labeling in complex samples such as eukaryotic cells[163]. I therefore sought to establish an alternative reference target that could be used for benchmarking PBSA approaches in cells to verify that improved trace quality for in situ data acquisition translated into an emitter counting performance with `quickPBSA` comparable to that observed in vitro.

Different intracellular targets have been proposed as references for quantitative fluorescence microscopy (see section 2.4). In addition to providing a structure with known stoichiometry, several practical aspects are important when choosing an in-cell reference target:

- Complexes should form with high efficiency and low or no dependence on protein expression level, since removal of incomplete complexes by purification is not possible when performing validation measurements in cells.
- Complex localization and subunit expression strength should exhibit minimal heterogeneity between cells.
- Complexes should exhibit intermediate densities to enable identification of individual complexes and sufficient sampling.
- Complexes should be sufficiently stable to remain assembled after fixation.

Two potentially suited reference targets which might meet the demands outlined above were recently suggested by Finan and colleagues and Thevathasan and colleagues. In the first report, Finan et al. demonstrated that enzyme homo-oligomers with 10-24 protein copies per complex can be used as in-cell standard for protein-counting applications[168]. In a second report, Thevathasan et al. proposed to use the eukaryotic nuclear pore complex as reference for benchmarking of super-resolution microscopy techniques and as calibration for quantitative SMLM[92].

To test which of these two in-cell reference complexes was better suited for benchmarking of emitter counting approaches based on PBSA, I first established and characterized the *E. coli* glutamine synthetase homo oligomer following the approach by Finan et al.. I then compared glutamine synthetase complexes prepared under optimized conditions to nuclear pore complexes containing labeled NUP107 as an alternative approach.

The structure of glutamine synthetase (GlnA) complexes has previously been determined by X-ray crystallography and was found to consist of 12 subunits assembled into two stacked hexameric rings[209]. In this configuration, individual Glutamine synthetase (GlnA) copies interact with at least three other subunits to form a stable complex with a total molecular weight of 12x51 kDa. In the previous work by Finan et al., GlnA has been used as standard in strongly disintegrated cells or after cell lysis[168]. Here, I wanted to find out, if GlnA can be used as standard in cells after stringent chemical fixation which aims at preserving cellular ultrastructure.

To address the need for establishing an robust in-cell standard with reproducible properties across cells and samples, I first addressed the question, whether GlnA complexes assemble in a reproducible fashion in a population of transiently transfected cells. For this, I transiently transfected HeLa cells with GlnA-HaloTag, and imaged chemically fixed samples by single-molecule sensitive widefield fluorescence microscopy.

I observed cells with widely different overall signal levels and diverging staining patterns (Fig. 4.22a). A subset of cells showed diffraction-limited signals as would be expected for GlnA clusters and was also observed by Finan et al.[168] (Fig. 4.22a, ’dim’). To determine if expression levels influenced the efficiency of GlnA-HaloTag complex formation, I used FACS to separate cells with different expression strengths. Also with flow cytometry, a large variability in expression strength across three orders of magnitude was observed. Most cells showed fluorescence signal above the level found in control cells (Fig. 4.22b). This indicated a high transfection efficiency and specific labeling of HaloTag with HTL-SiR was achieved. I used FACS to isolate four fractions of transfected cells with overall signals higher than the signal encountered in untransfected cells and gradually increasing expression strength (assuming SiR signal correlated with expression strength). From each fraction, I prepared samples for fluorescence microscopy immediately after cell sorting and manually scored a random selection of cells according to the observed staining pattern (Fig. 4.22c). Here, the percentage of cells showing formation of structures with mainly diffraction-limited signals (equivalent to cells labeled as ’good’ in Fig. 4.22a) did not exceed 40 % for any of the sorted fractions. The largest fraction of cells with mainly diffraction-limited signals was observed in fractions I and II with low overall HTL-SiR signals during sorting (Fig. 4.22c). In general, increased HTL-SiR signal was correlated with the percentage of cells showing continuous signal without distinguishable diffraction-limited features. These results indicate that additional factors beyond the transgene expression strength in a given cell contribute to observed staining patterns.

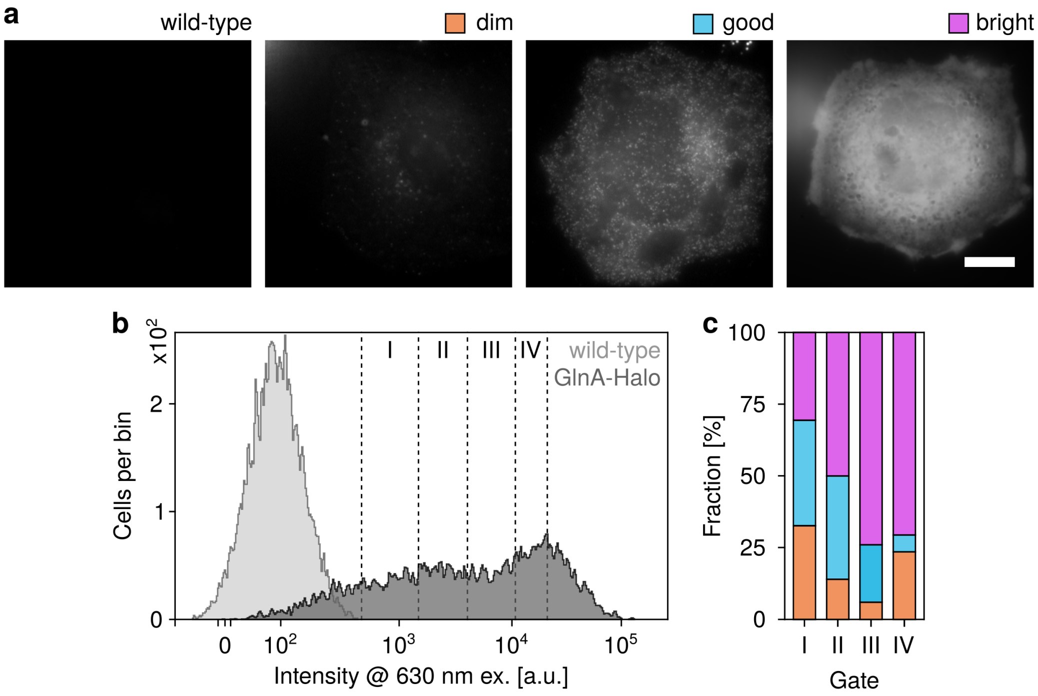

Figure 4.22 GlnA-HaloTag expression level and phenotypes vary in transiently transfected HeLa cells. HeLa cells transiently transfected with GlnA-HaloTag or HeLa wild-type cells were labeled with 10 nm HTL-SiR for 60 min, fixed with 3.7 % PFA and imaged by widefield fluorescence microscopy or subjected to FACS without fixation. **a,** Representative cells exhibiting different GlnA-HaloTag expression strength and HeLa wild-type cells. Scale bar: 10 µm. **b,** FACS analysis of fluorescence intensity per cell for HeLa cells transiently expressing GlnA-HaloTag (dark gray) or wild-type HeLa cells (light gray). For each cell type, 1×10^4 cells were processed and sorted. Vertical dashed lines indicate FACS gates of fractions I-IV. **c,** Percentage of cells exhibiting low, good or bright phenotypes in widefield imaging for each FACS fraction. 50 cells were assessed for fractions 1-3. Fraction 4 contained a large number of non-adherent cells, only 17 cells in total were assessed for this fraction. Color coding as in a.

To further evaluate the suitability of GlnA-HaloTag as reference target, I reasoned that stable integration of GlnA-HaloTag into the host cell genome might help to obtain a cell population with stable expression of GlnA-HaloTag at intermediate expression levels.

With the help of Florian Salopiata (DKFZ Heidelberg), we therefore created a stable HeLa GlnA-HaloTag cell line via viral transduction using the Phoenix Ampho system[173]. For this purpose, GlnA-HaloTag was transferred from its original plasmid where the gene was expressed under control of a CMV promoter to the pBabe plasmid which was also used as backbone for viral transduction of gSEP constructs in Section 3.2.1. After transduction, stably transduced cells were selected using $6.0\,\mu g\,mL^{-1}$ puromycin as selection antibiotic and a population of cells surviving under selection was obtained two weeks after transduction (Fig. B.6).

To assess the effect of selection strength with varying concentrations of puromycin on GlnA-HaloTag expression, I compared selected different cell populations with varying puromycin concentration and compared the observed GlnA-HaloTag signals by widefield fluorescence microscopy. Here, I observed considerable cell-to-cell variation in signal distributions as previously observed after transient transfection and FACS for all tested puromycin concentrations. No differences for cells selected under $1.5\,\mu g\,mL^{-1}$, $3\,\mu g\,mL^{-1}$ and $6\,\mu g\,mL^{-1}$ puromycin with respect to the median cluster intensity or cluster number per cell were observed (Fig. B.6). The cell population grown under puromycin selection at $6\,\mu g\,mL^{-1}$ was used for further experiments to exert maximum selection pressure on cells.

Flow cytometry of stably transduced cells after two weeks of puromycin selection again showed large variability in signal after staining with HTL-SiR indicating varying GlnA-HaloTag expression levels (Fig. 4.23a). In contrast to transiently transfected cells directly after transfection, the stably transduced cell population showed two peaks in HTL-SiR signal, one at low signal strength and a second peak at higher signal strength.

Analogous to experiments performed after transient transfection, I sorted HeLa GlnA-HaloTag cells into four fractions by FACS and imaged them by widefield fluorescence microscopy after fixation. In all fractions, cells with bright, diffraction-limited punctae were observed (Fig. 4.23c). In contrast to staining patterns observed after transient transfection, fraction I of stably transduced HeLa GlnA-HaloTag cells showed very few and less bright signals. In fraction IV, cells tended to exhibit increased background signal. The two intermediate FACS fractions (termed ’low’ and ’high’) showed largest proportion of cells with bright punctae and were used for further experiments. In the course of further experiments, I noticed that the expression of GlnA-HaloTag over time seemed unstable. While immediately after sorting, a substantial fraction of cells showed high intensity GlnA-HaloTag clusters, the fraction of cells containing clusters decreased with increasing age of the respective culture. For example, a culture of unsorted HeLa GlnA-HaloTag cells showed a

45 % median decrease in the number of GlnA-HaloTag clusters per cell within 30 days under identical staining and fixation conditions (Fig. 4.23b). This finding was further confirmed by FACS analysis, where I re-assessed the cell-wise HTL-SiR signal levels observed in the two intermediate fractions ’low’ and ’high’ 24 days after initial cell sorting (Fig. 4.23d,e). Within this time, the observed cell-to-cell signal levels broadened considerably beyond the gates applied during initial sorting and only a mild enrichment of $\sim$1.6x over the unsorted population was achieved (Fig. 4.23f).

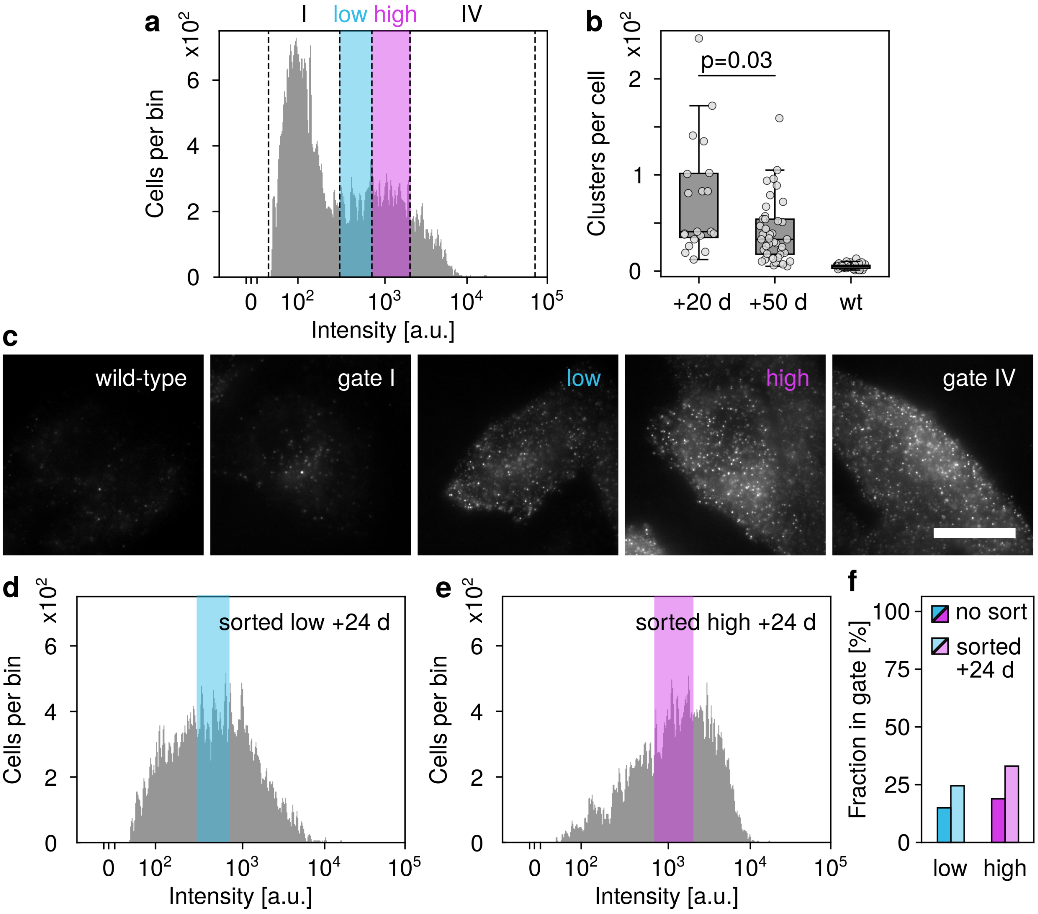

Figure 4.23 Stability of GlnA-HaloTag expression in stably transduced HeLa cells over time. Cells were labeled with 10 nm HTL-SiR for 60 min and subjected to flow cytometry without prior fixation or to widefield fluorescence microscopy after fixation with 3.7 % PFA. **a,** Stably transduced cells were assessed by flow cytometry and sorted into indicated fractions by FACS. **b,** The number of GlnA clusters per cell was measured by widefield fluorescence microscopy and SMLM analysis. HeLa cells stably expressing GlnA-HaloTag were labeled, fixed and imaged at different time points after initial transduction. As reference, wild-type (wt) cells were labeled and imaged under identical conditions. Samples were compared using Welch's *t*-test. **c,** Widefield fluorescence microscopy images of representative cells from FACS fractions shown in b and HeLa wild-type cells labeled under identical conditions. **d,e,** FACS analysis of HeLa GlnA-HaloTag cells from low (d) and high (e) gates cultured for 24 days after sorting. Gates applied during initial sorting highlighted in corresponding colors. **f,** Fraction of cells within respective gates for cultures sorted 24 days prior to re-sorting (light colors) and without prior sorting (dark colors) for low and high expressing fraction.

From these experiments, I concluded that while stable transduction enabled simplified experiments without the need for repeated transfection, the stable expression levels and low cell-to-cell heterogeneity in cluster density desired for a robust standard target were difficult to achieve using GlnA-HaloTag complexes as target.

Despite the varying expression strength of GlnA-HaloTag after stable transduction, a subset of cells exhibited expression levels as would be required for use as in-cell standard. I therefore decided to further evaluate the behavior of GlnA-HaloTag under varying fixation conditions with the aim to identify sample preparation conditions which minimized background signal while at the same time avoiding strong disintegration of cells during fixation. For this, stably transduced and puromycin-selected (6.0 $\mu g\,mL^{-1}$) HeLa GlnA-HaloTag cells were labeled with HTL-SiR (10 nM, 60 min) and fixed with three different fixation buffers. The first buffer contained 3.7 % PFA as fixative and was identical to the fixation used for previous experiments with GlnA-HaloTag shown above. In addition, I tested a high stringency fixation which included addition of GA. The third fixation approach followed previous experiments by Finan et al. and included a post-fixation extraction step with 0.5 % TX-100to remove GlnA-HaloTag not immobilized during fixation.

Widefield fluorescence microscopy showed that the overall intensity of cells, the number of visible clusters and cluster intensities seemed to depend on the specific fixation protocol (Fig. 4.24a). For a quantitative comparison of GlnA-HaloTag cluster properties under different fixation conditions, I performed automated imaging of ~50 cells per condition and assessed the number and intensity of individual cluster by sub-pixel localization. GlnA-HaloTag cluster intensities showed a high variability for each of the tested fixation conditions with >20-fold differences in cluster intensity within a single condition (Fig. 4.24b). Stringent fixation with a combination of PFA and GA resulted in the highest median cluster intensity almost 2-fold higher than fixation with PFA alone and a tendency towards high intensity clusters.

On a per-cell basis, I observed that for all three fixation conditions, a range of different cluster numbers per cell (100-500) could be observed and that also the median intensity of clusters in single cells varied between cells (Fig. 4.24c). At the same time, the number of clusters per cells and the mean intensity of clusters per cell varied substantially within each of the three fixation conditions (Fig. 4.24). Cluster numbers varied from 100-500 per cell. Also the mean cluster intensity per cell showed ~500 % spread. On average, stringent fixation with PFA supplemented with GA resulted in increased median cell-wise cluster intensities. An additional extraction step after fixation resulted in an increase in detected clusters per cell. For all fixation conditions, the overall properties of HeLa GlnA-HaloTag cells with respect to signal intensities and the number of detected clusters were different from wild-type cells subjected to the same labeling and fixation conditions for the majority of cells. Only cells exposed to extraction with TX-100 post fixation, in the number of transfected cells that were indistinguishable from wild-type cells (Fig. 4.24d).

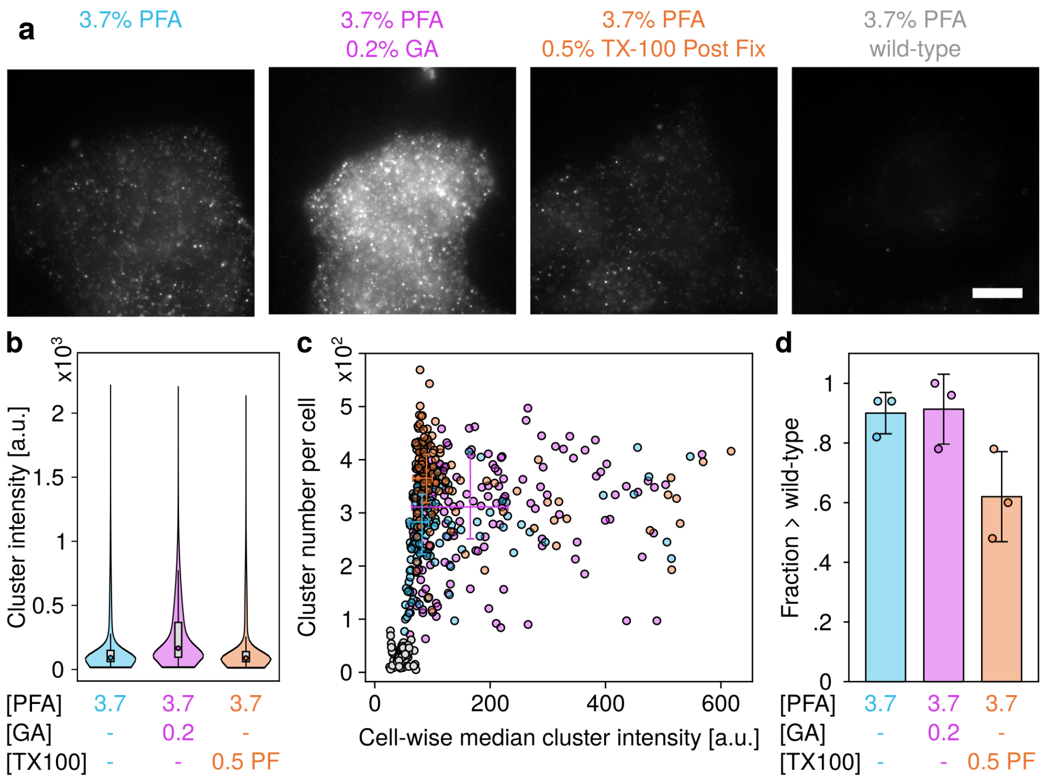

Figure 4.24 Influence of fixation condition on GlnA-Halo cluster intensity and number. Stably transduced HeLa GlnA-HaloTag cells were labeled with 10 nm HTL-SiR for 60 min and chemically fixed in indicated fixation buffers containing PFA or PFA and GA. For indicated conditions, samples were permeabilized with 0.5 % TX-100 after fixation. Samples were imaged by widefield fluorescence microscopy and cluster properties were obtained by subsequent SMLM analysis. **a,** Representative images from stably transduced HeLa GlnA-HaloTag cells for indicated fixation conditions and HeLa wild-type cells with 3.7 % PFA fixation. Scale bar: 10 µm. **b,** Cluster intensities across all clusters identified in all imaged cells for the indicated conditions. **c,** Cell-wise cluster number and median cluster intensity for different fixation conditions (condition indicated by same colors as in a). Gray data points represent wild-type cells across under identical labeling conditions and all fixation conditions. For each fixation condition, a two-axis error plot indicates mean±SD cluster intensities and cluster number per cell. **d,** Fraction of cells in HeLa GlnA-HaloTag samples with median cluster intensity and cluster number >95 percentile of corresponding wild-type sample under identical fixation conditions. Bar represents mean across three samples, error bar represents SD.

I additionally investigated the influence of extraction with TX-100 during fixation by performing an equivalent analysis of cluster number and cluster intensity in dependence on TX-100 concentration. For all tested TX-100 concentrations (0.02 %, 0.1 % and 0.2 %), I observed a decrease in the cluster number per cell and a dependence of cluster intensity on TX-100 concentration.

Based on these experiments, I concluded that while it was possible that GlnA-HaloTag could be used as reference target under selected sample preparation conditions and after further optimization of GlnA-HaloTag expression, its use as universal reference standard is limited at the current stage. If the established HeLa GlnA-HaloTag cell line is to be used as reference standard in the future, the most promising experimental strategy would be to minimize the time after cell sorting to maximize the fraction of cells in a culture with suited expression levels. This cell population should then be fixed with PFA supplemented with TX-100 at low concentrations followed by manual selection of cells for imaging.

As GlnA-HaloTag had therefore been shown to be unsuited as reference standard at this stage, I then tested if a genome-edited U2OS cell line expressing NUP107-SNAP$_f$-tag (named NUP107-SNAP hereafter) with a known copy number of 32 per NPC could be used as alternative reference standard[210]. As recently suggested by others, the nuclear pore complex is a promising candidate for use as reference standard since it has several desirable features: The NPC is an essential and endogenous complex in eukaryotic cells. Further the NPC has an symmetric structure with conserved stoichiometry and its integration into the nuclear membrane should render it resistant against aggregation during fixation[92,155].

Labeling of U2OS NUP107-SNAP with BG-SiR followed by chemical fixation and widefield fluorescence imaging revealed the expected nuclear labeling. Due to the stable genomic integration of the SNAP$_f$-tag fusion protein, consistent staining of nuclei across many cells was observed. At the equatorial plane of cells, labeling of NUP107-SNAP resulted in individual pores being visible at the rim of the nucleus (Fig. 4.25a), while imaging at the basal nuclear membrane revealed individual punctae (Fig. 4.25b). In both cases, negligible cytosolic background signal was observed. Cross-sections through nuclei at both axial planes further showed that punctae appeared diffraction-limited and that the intensities of individual NPCs showed limited variability (Fig. 4.25c). To evaluate the influence of fixation on the number of detected NPCs per nucleus and on the intensities of individual NUP107-SNAP, I compared two fixation protocols with and without pre-extraction using TX-100 using the same SMLM approach used for characterization of GlnA-HaloTag above. For both fixation protocols, the number and median intensity of clusters in individual cells showed a low cell-to-cell variations of 10-20 %. However, pre-extraction with 0.5 % TX-100 prior to chemical fixation with 2.4 % PFA as suggested by Thevathasan et al.[92], resulted in a reduced detected cluster numbers per cell as compared to a fixation protocol without pre-extraction (Fig. 4.25d). No differences in the median cell-wise cluster intensity between both fixation protocols were observed (Fig. 4.25e).

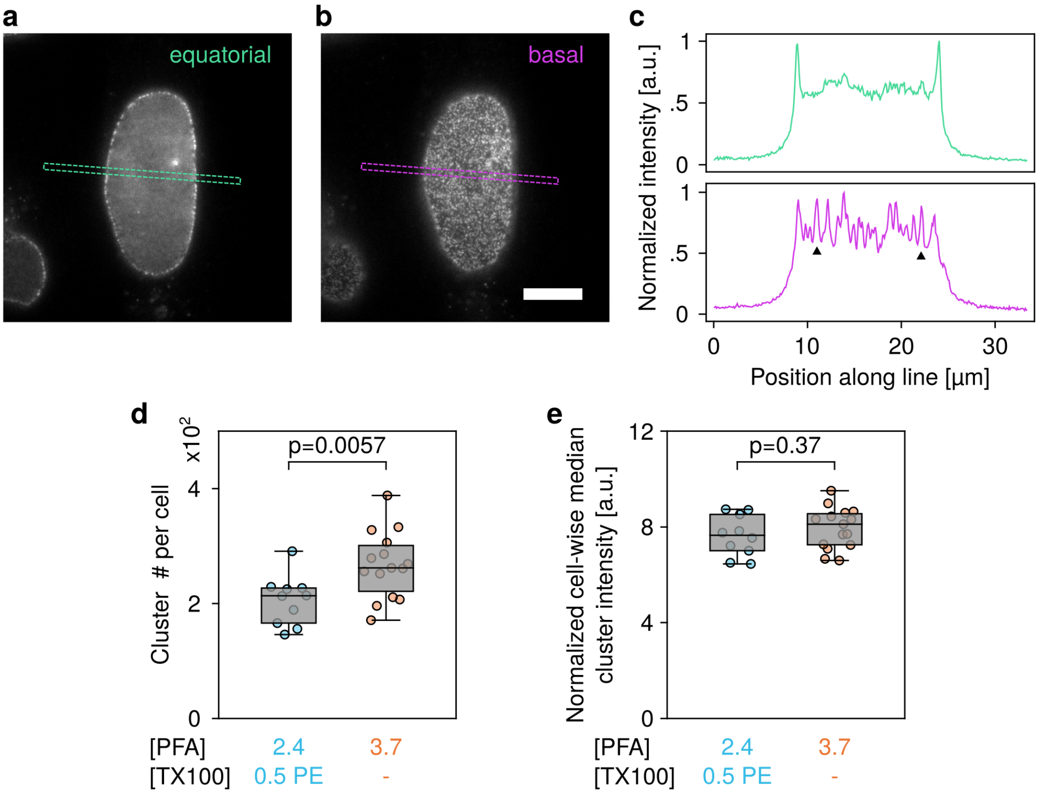

Figure 4.25 Evaluation of NPCs with labeled NUP107-SNAP as standard target. U2OS cells stably expressing NUP107-SNAP were labeled with 200 nM BG-SiR for 150 min (a-c) or 120 min (d-f), washed and fixed with 3.7 % PFA or pre-extracted with 0.4 % Triton X-100 and fixed with 2.4 % PFA. Cells were imaged by widefield fluorescence microscopy in PBS. **a,** Equatorial section through nucleus. **b,** Basal section through nucleus. Scale bar a,b: 10 µm. **c,** Intensity profiles along boxes indicated in a,b. Arrow heads indicate putative positions of individual nuclear pore complexes in profile across basal nuclear membrane. Profiles were normalized to their respective peak intensity. **d,e** Influence of sample fixation on cluster number (d) and median cluster intensity per cell (e) in images of basal nuclear membrane. Circles: results from individual cells. Cluster intensities were normalized to the respective localization offset to correct for differences in acquisition settings. Statistical comparisons were performed using Welch's *t*-test.

As final comparison between both candidate reference targets, GlnA-HaloTag and NUP107-SNAP, I compared the detected cluster density and the variability of cluster intensities within individual cells. Both parameters are important since they inform about how many cells will have to be imaged for subsequent PBSA analysis and about the homogeneity of complex stoichiometry within cells.

To compare the density of detected clusters per cell, I segmented whole cells (GlnA-HaloTag) or the nuclei of cells (NUP107-SNAP) via manual tracing (GlnA-HaloTag) or intensity-based thresholding (NUP107-SNAP). Different segmentation approaches were required since for GlnA-HaloTag strong variation of cluster density and cluster intensity complicated identifying robust segmentation algorithm. As noted previously, the density of GlnA-HaloTag clusters in HeLa cells stably expressing GlnA-HaloTag depended strongly on fixation condition. After fixation with PFA or a combination of PFA and GA, median cluster densities around $0.2\,\mu m^{-2}$ were observed, while extraction with TX-100 post fixation resulted in reduced cluster densities typically around $0.1\,\mu m^{-2}$. To a lesser degree, this was also the case for the U2OS NUP cell line (Fig. 4.26a). Here, the median density of detected clusters was around $0.7\,\mu m^{-2}$ after fixation with PFA and around $0.6\,\mu m^{-2}$ when pre-extraction with TX-100 prior to PFA fixation was performed. The relative density variability between cells from the same condition was slightly higher for GlnA-HaloTag and the overall density for NUP107-SNAP was 2-4-fold higher than the density of detected clusters in GlnA-HaloTag samples. This is important since a high number of clusters per cell allows for statistically robust sampling by imaging a small number of cells. At the same time, this also increases the chance for overlapping PSFs from neighboring clusters.

As final parameter used to compare the behavior of both reference targets, I determined the variability in cluster intensity. This metric is especially important since it allows to assess possible variations in cluster stoichiometry. As measure for cluster intensity variations within and across cells, I made use of the Quartile coefficient of dispersion (QCD), which is defined as:

$$QCD = \frac{Q_3 - Q_1}{Q_3 + Q_1} \tag{4.1}$$

Where Q_1 and Q_3 are the first and third quartiles of a given distribution[211]. A lower QCD indicates less variation within an individual cell, while the spread of QCD values across cells for a given condition is a measure for cell-to-cell variation. Using this metric, I observed that GlnA-HaloTag showed substantial QCD variability across cells within individual fixation conditions and differences between fixation conditions. NUP107-SNAP showed overall lower QCD and very low QCD variability between cells (Fig. 4.26b).

While GlnA was used as reference standard under non-native cell sample preparation conditions before[168], the results described above show that using GlnA-HaloTag as reference standard in intact cells is challenging due to the strong influence of fixation conditions on complex density and intensity. Furthermore, the unstable expression of the established HeLa GlnA-HaloTag cell line over time and the high cell-to-cell variation in expression strength would have to be addressed further before routine use as reference standard. In contrast, NUP107-SNAP appeared to be more suited as reference standard since the cell-to-cell variability in cluster density and cluster intensity was substantially lower than observed for GlnA-HaloTag. In addition, the cluster intensity was not influenced by the applied fixation protocol. The fact that NUP107-SNAP was introduced into U2OS via genome editing also suggests that the expression strength over time for NUP107-SNAP should be constant. One disadvantage of NUP107-SNAP as reference standard is the relatively high density of NPCs on the nuclear membrane which could lead to interfering signal form individual NPCs located in close proximity. While this potentially complicates subsequent PBSA data analysis, an adapted candidate ROI identification and trace extraction approach could help to reduce the impact of partially overlapping PSFs from neighboring clusters.

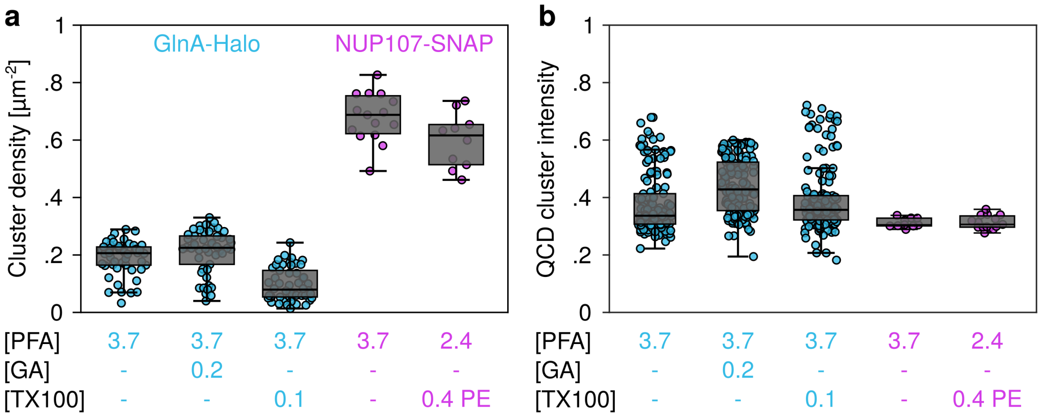

Figure 4.26 Comparison of GlnA-HaloTag and NUP107-SNAP as standard targets. HeLa cells stably expressing GlnA-HaloTag or U2OS cells stably expressing NUP107-SNAP were labeled with 10 nM HTL-SiR or 200 nM BG-SiR, fixed as indicated and imaged by widefield fluorescence microscopy in PBS. **a,** Comparison of cluster density in HeLa GlnA-HaloTag and U2OS NUP107-SNAP for indicated fixation conditions. GlnA-HaloTag density was computed using manually determined cell outlines, NUP107-SNAP density was computed using automatically determined nuclei outlines. **b,** QCD of cluster intensity in HeLa GlnA-HaloTag and U2OS NUP107-SNAP for indicated fixation conditions. a,b Circles: results from individual cells.

4.5.3 Extended workflow is required for in situ counting

After having established that NUP107-SNAP could be used as in-cell reference standard for validating the `quickPBSA`framework, I set out to adapt the `quickPBSA` framework previously validated in vitro to the more complex image data obtained from in-cell measurements (Fig. 4.27). As shown in section 4.4 (e.g. Fig. 4.8b), localization of diffraction-limited signals in images acquired on in-vitro samples was straightforward. In contrast ROI detection and trace extraction on image stack acquired on cellular samples is considerably more complex as it is not possible to immobilize target structures at low density and on a planar surface. Instead, trace extraction in cellular samples has to be able to deal with high emitter densities, partially overlapping PSFs and elevated background signal due to fluorophores located outside the focal plane. To enable analysis of cellular samples, the candidate ROI filtering module was included for selecting suited ROIs based on the localization properties of individual signals and the local density of localized signals (Fig. 4.27b). This is critical since due to the variable density of signals encountered in cells, erroneous localizations and localizations from non-diffraction limited signals are more likely to occur than with data from in vitro samples. The candidate filtering module also serves to remove candidate ROIs located in close proximity to other ROIs from further analysis. In addition, the trace extraction module was extended to enable local background correction of traces. This is crucial since due to the three-dimensional structure of cellular samples, signal from complexes in other axial positions may contribute to signal extracted from a given ROI (Fig. 4.27c).

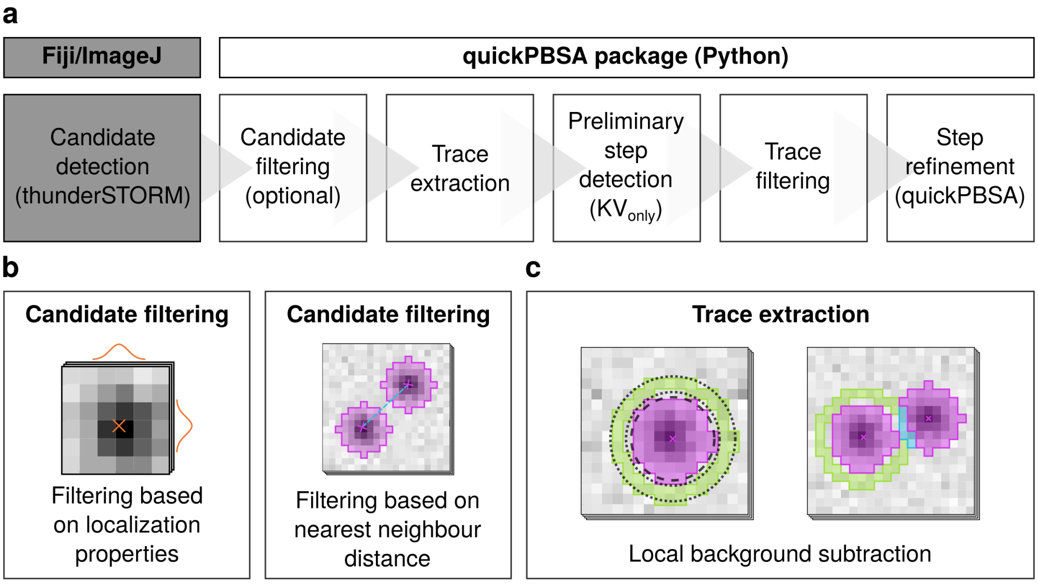

Figure 4.27 Extended quickPBSA workflow for in situ PBSA. a, Extended `quickPBSA` workflow with active candidate filtering module and expanded trace extraction module. **b,** Candidate filtering based on sub-pixel localization properties and distance between centroids of neighboring ROIs. **c,** Left: Principle of local background subtraction using ring-shaped region (green) with inner radius r_{bg1} and outer radius r_{bg2} around candidate ROI (magenta) shown as dashed lines. Right: Local background subtraction after removal of regions (blue) overlapping with neighboring candidate ROIs (magenta). Final background region shown in green.

During manual inspection of results from SMLM-based candidate ROI detection with `thunderSTORM`, I observed that unlike for data recorded from in vitro samples, ROI detection was less reliable for data recorded from NUP107-SNAP expressing U2OS cells. While in general, SMLM signal detection was sufficiently sensitive to detect the majority of putative NPC signals, increased background and overlapping signal led to spurious ROI detection events (Fig. 4.28a-c). Since sub-pixel localization of signals provides additional information beyond the centroid coordinates, I decided to make use of these information for removing spurious ROI detection events. NPCs have a diameter of 120 nm[6], which is below the diffraction limited of typical light microscopes. Thus, the width of the fitted 2D Gaussian function σ can help to identify diffraction-limited signals and to distinguish them from signals originating from partially overlapping diffraction-limited signals or signals from structures outside the focal plane[2]. Setting a lower limit for σ can help removing spurious ROI detection events originating from noise rather than diffraction-limited signals. On NUP107-SNAP, candidate ROI filtering based on σ resulted in removal of ~25 % of

---

[2] See also Section 3.4.3 where I applied a similar procedure in the context of labeling efficiency measurements with gSEP.

initially detected candidate ROIs (Fig. 4.28d), which mainly localized to background regions (Fig. 4.28b,c).

The random arrangement of complexes within cells will inevitably lead to a fraction of complexes being localized in close proximity to each other. For these complexes, extraction of intensity traces via a ROI with fixed size is less reliable since signal from one complex may contribute to signal in ROI from another complex. To remove complexes for which such effects could not be excluded, I introduced a spatial filtering routine. This routine removes ROIs for which the centroid is located closer than a user-defined threshold to the centroid of neighboring ROIs (Fig. 4.27b). The relatively high density of signals in U2OS NUP107-SNAP cells, resulted in removal of a majority of candidate ROIs with only $\sim$20 % of ROIs remaining after spatial filtering (Fig. 4.28d). While candidate filtering has a profound effect on the number of NPCs suited for PBSA, we reasoned that including a spatial filtering step would increase the quality of photobleaching traces substantially and was therefore necessary for robust emitter counting with `quickPBSA`. It is, however, also important to note that despite applying the filtering routines described above, it can not be excluded that individual ROI contain multiple complexes if they are located at a proximity less than the diffraction-limit of the microscope used.

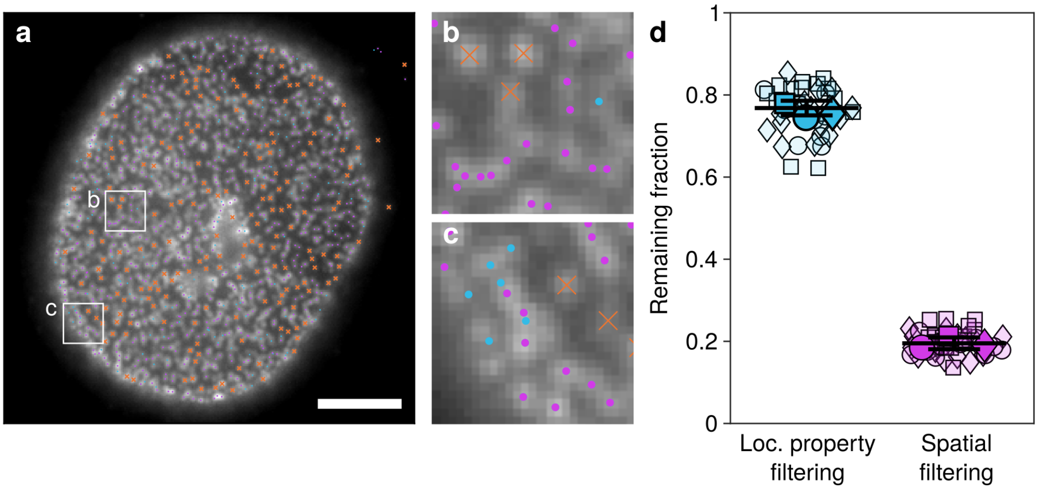

Figure 4.28 Filtering of candidate positions based on localization properties. U2OS cells stably expressing NUP107-SNAP were labeled with HTL-SiR, fixed and imaged in ROXS PCD buffer by widefield fluorescence microscopy on setup A. Candidate positions for PBSA were identified by SMLM analysis. **a,** Overview image of representative cell with overlayed candidate positions (colored symbols). Candidates removed based on their localization properties (sigma 100-200 nm, cyan) and nearest-neighbor distance (<475 nm, magenta) indicated by dots. Positions subjected to downstream trace extraction highlighted by orange crosses. Scale bar: 5 µm. **b,c,** Magnified images of boxed regions in a. **d,** Fraction of candidate spots remaining after corresponding filtering step. Small symbols: Results from individual cells, large symbols mean from experiment. Error bar: Mean±SD across all means from experiments.

The background signal visible between individual diffraction-limited signals in cellular samples was substantially higher than background signals typically observed on in vitro samples. From image data acquired on in vitro samples, photobleaching traces with high SNR could be obtained by extracting frame-wise intensities from ROIs centered on diffraction-limited signals within images without additional processing. In cells, this approach frequently resulted in photobleaching traces without discernible step-like intensity changes even after spatial filtering of candidate ROIs.

Traces can be corrected for local background by subtracting the frame-wise intensity of a region in close proximity to the target ROI. Here, it is important to make sure that the background region does not overlap with the PSF of the target itself to ensure that no signal emitted by fluorophores contained in the target complex contribute to the background signal (illustrated in Fig. 4.27c, left). In cells, the high density of candidate ROIs usually does not allow for blind extraction of background intensities from ring-shaped regions as it was previously done for data acquired on in-vitro samples (Fig. 4.29a,b). I therefore implemented an adaptive local background correction routine, which takes detected neighboring ROIs into account and removes parts of the background region overlapping with other candidate ROIs (Fig. 4.27c, right). This procedure performed robustly on data recorded on NUP107-SNAP U2OS cells and produced background ROIs which typically did not overlap with other diffraction-limited signals (Fig. 4.29c). This adaptive local background correction was found to be especially important since the intensity decays of background ROIs and candidate ROIs were indistinguishable, suggesting that the main background contribution originates from out-of-focus complexes labeled with the same fluorophore (Fig. 4.29d,e).

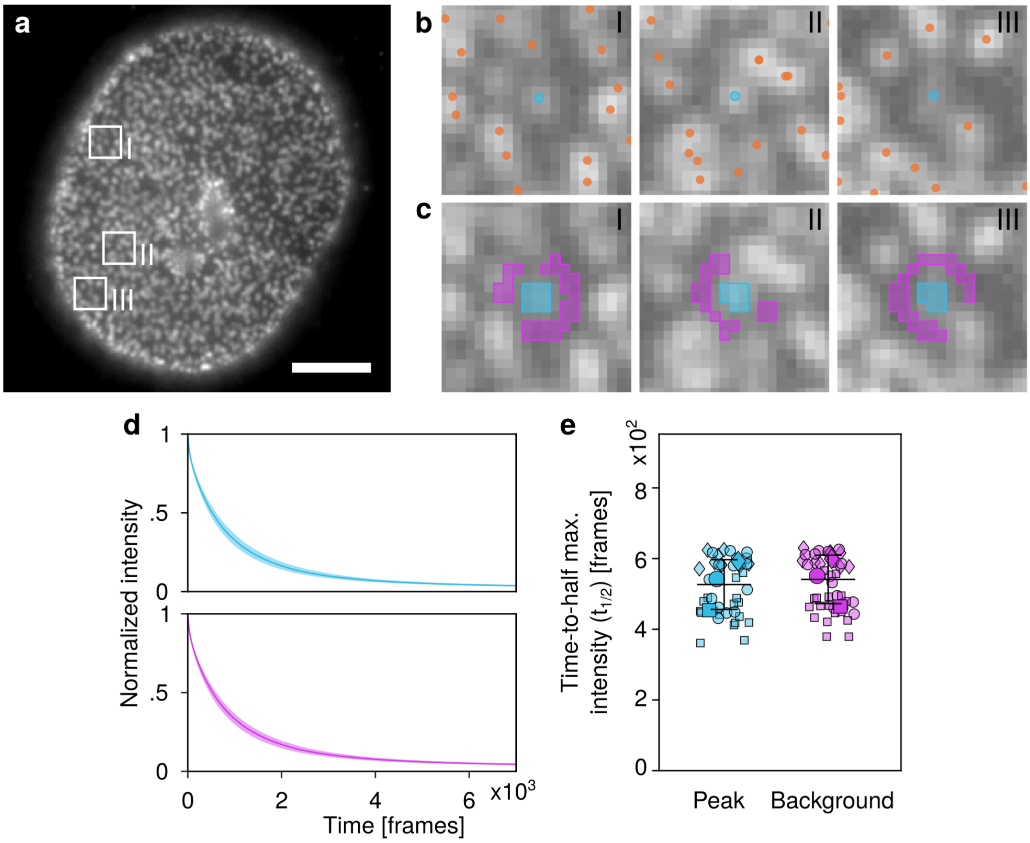

Figure 4.29 Background subtraction for in situ photobleaching trace extraction. a, Scale bar: 5 µm. **b,** Magnified views of regions indicated in (a). Central ROI (cyan) and neighboring candidate positions (orange). **c,** ROI (cyan) and background region (magenta) ROIs for peak ROI positions shown above. **d,** FOV-wise averaged photobleaching curves for peak (cyan, top) and background (magenta, bottom) ROIs. **e,** Decay times to half maximum intensity ($t_{1/2}$) for FOV-wise averaged peak and background ROIs. Large symbols: mean $t_{1/2}$ per experiment, small symbols: individual FOVs. Error bars indicate mean and SD across experiment means.

4.5.4 NUP107 copy number in human nuclear pore complexes

In the course of the last section I showed how different aspects of an PBSA experiment such as the choice of fluorophores or buffer systems and routines for extraction of suited intensity traces from densely-labeled cellular samples can be optimized to improve input data quality. In the final experiments presented in this chapter, I will make use of this optimized pipeline for data acquisition to demonstrate the performance of `quickPBSA` for emitter counting by PBSA in situ. As validation target for these measurements, I used the gene-edited U2OS NUP107-SNAP cell line characterized above and labeled SNAP$_f$-tag with BG-SiR. Data was then recorded in ROXS PCD buffer. Acquired data was processed with the full `quickPBSA` framework including the candidate filtering module and trace extraction with local adaptive background correction.

In raw intensity traces extracted from the recorded dataset without background correction, virtually no step-like intensity changes could be observed visible (Fig. 4.30a,b). After local adaptive background subtraction, traces exhibited step-like intensity changes and plateaus between changes (Fig. 4.30c) indicating that the photostability was sufficient to record intensity traces amenable for processing by PBSA. However, the SNR of recovered traces was substantially lower than for previously analyzed traces from in vitro measurements. As first step to verify that trace interpretation with `quickPBSA` produced trace models reflecting likely photobleaching events within traces, I first manually selected a small subset of traces where individual steps were discernible and compared these traces to the corresponding `quickPBSA` model. To facilitate manual comparison between the trace models and raw data, I adjusted the model trace using a linear scaling factor to minimize differences between raw data and model (Section 6.3.5). As shown in Fig. 4.30d,e, `quickPBSA` indeed produced models, which matched visible features of the underlying raw data trace including photobleaching steps and transient dark states of individual fluorophores. In contrast to this, step detection with the KV$_{\text{only}}$ algorithm typically failed to produce reasonable trace models (Fig. 4.30f).

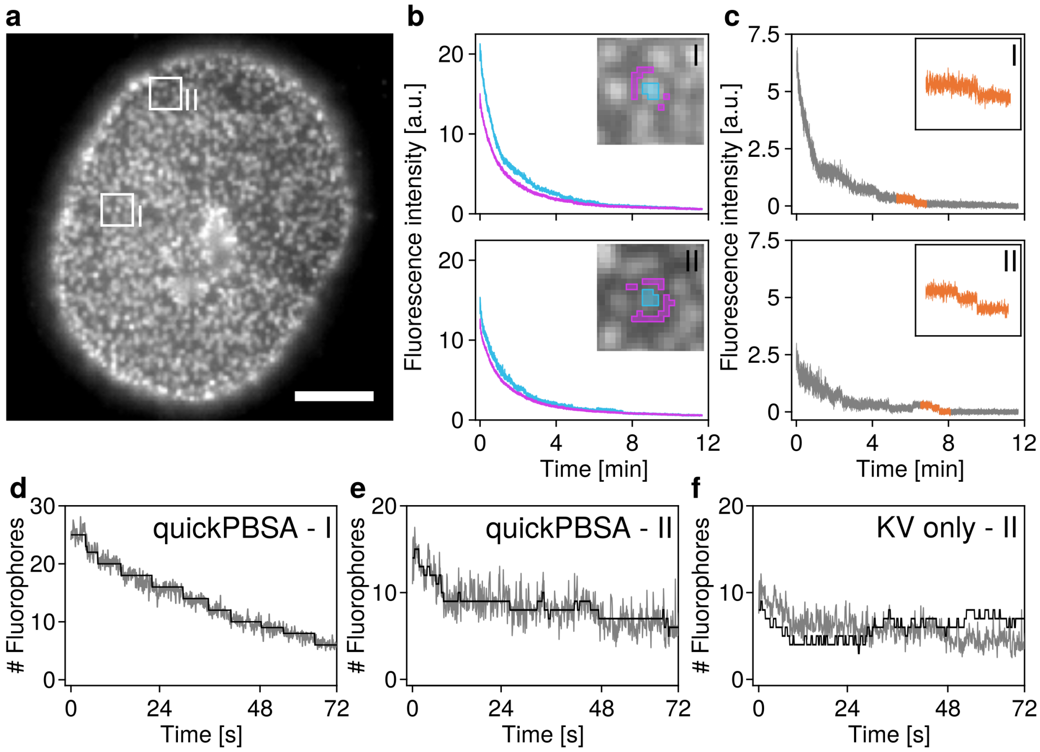

Figure 4.30 Optimized trace extraction enables extracting traces suitable for quickPBSA processing. U2OS cells stably expressing NUP107-SNAP were labeled with 200 nm BG-SiR for 120 min and fixed with 3.7 % glsPFA as described in Section 6.1.5. Samples were then imaged on widefield setup A. Identification of candidate trace positions, trace extraction and trace analysis was performed as described in Section 6.3.5. **a,** Overview image of representative cell exhibiting nuclear localization of NUP107-SNAP and diffraction-limited features. Scale bar: 5 µm. **b,** Photobleaching traces of peak (cyan) and background (magenta) ROI for individual peak positions. Inset: Zoom-ins of regions indicated in a and used for trace extraction. **c,** Background-corrected photobleaching traces obtained from region indicated in a,b exhibit expected step-wise photobleaching. Zoom-ins: time windows around ultimate photobleaching step highlighted in orange. **d-f,** Zoom-ins of individual photobleaching traces from ROIs highlighted in a,b (gray) with corresponding result of emitter number estimation with indicated algorithm (black). Photobleaching traces were smoothed with a five point running mean filter for visual inspection.

To test the performance of `quickPBSA` on a broader scale, I proceeded by performing a full analysis on a larger dataset. For this, I acquired time lapse images (7×10^3 per image set) on U2OS NUP107-SNAP cells prepared as described above to produce a dataset consisting of 32 cells in two independent experiments. From this data, 6.3×10^3 traces after candidate ROI filtering were obtained and subjected to step detection with `quickPBSA`. On average, 55 ± 9 % of traces were successfully processed (Fig. 4.31a), which was ~20 % below the success rate achieved for analysis of recorded on DNA origami. The average processing time per trace was 100 ± 5 s (Fig. 4.31b), about 10-fold higher than for the benchmarking dataset based on semi-synthetic data (3×10^3 data points per trace vs. 7×10^3 data points per trace here). While the number of traces extracted from individual cells varied by more than four (59-204 traces per cell overall), the average number of detected photobleaching steps per NPC was relatively constant across cells (21.3 ± 0.9, mean±SD) indicating that $\sim$60 successfully analyzed traces were sufficient to determine the mean stoichiometry of NUP107-SNAP in a cell (Fig. 4.31c). In line with the results from individual cells, the distribution of detected photobleaching steps per NPC across all cells exhibited a unimodal shape with a mean step number of 21.3 ± 7.6 (mean±SD) (Fig. 4.31d). In contrast, the local intensity approach for estimating emitter numbers initially present in NPC prior to photobleaching produced a strongly skewed distribution with two visible peaks and a mean of 27.1 ± 13.8 (Fig. 4.31e). Since such behavior could be an indicator for the presence of two NPCs in the extracted ROIs, I modeled both distributions with a coupled double-Gaussian function (see Section 6.3.5). For `quickPBSA`, this resulted in a predicted mean fluorophore number of 10.9 ± 0.2 (95 % CI) and a standard deviation of 3.8 ± 0.3 fluorophores per NPC. The fraction of ROIs predicted to contain contributions from two NPCs was 83 %. The fit to the distribution obtained with the local intensity approach resulted in a mean fluorophore number of 15.7 ± 0.7, a standard deviation of 4.8 ± 0.4 per NPC and a fraction of 45 % ROIs predicted to contain contributions from two NPCs.

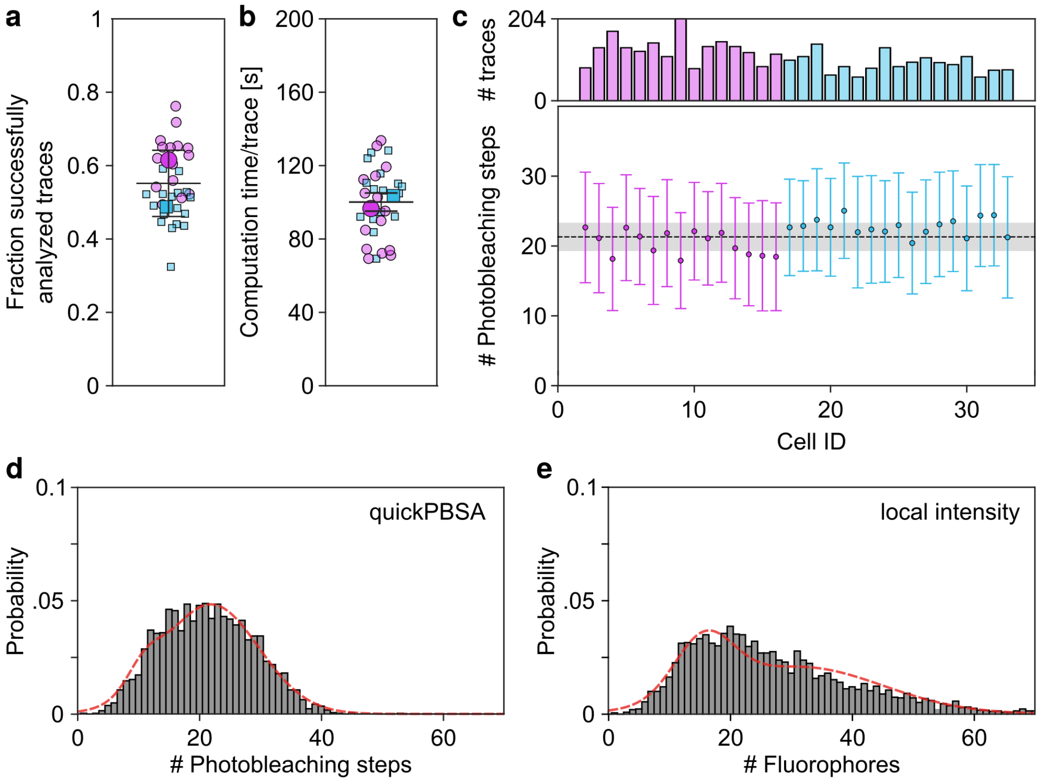

Figure 4.31 quickPBSA provides robust NUP107-SNAP copy numbers in situ. NUP107-SNAP in gene-edited U2OS cells was stained 10 nM HTL-SiR for 120 min, fixed in 3.7 % PFA and imaged in ROXS PCD buffer. Two independent experiments were performed with one experiment shown in cyan and one experiment shown in magenta. **a** Fraction of successfully analyzed traces per cell. Large symbols correspond to experiment mean. Median±SD across experiment means indicated by error bar. **b,** Cell-wise mean computation time per trace. Symbols as in (a). **c,** Top: Number of successfully analyzed traces per cell. Bottom: Mean±SD (circles, error bars) photobleaching steps per NPC per cell from corresponding cell in plot above. Dashed line and shaded region indicate mean±SD across all NPCs from all cells. **d,** Distribution of photobleaching steps per trace detected with `quickPBSA` in successfully analyzed traces (3.5×10^3) pooled from two experiments. Red lines: linked double-Gaussian model fitted to distribution. R^2_{adj}: 0.98. **e,** Estimated fluorophore number from local intensity approach on identical set of traces processed with `quickPBSA` in (d). R^2_{adj} for linked double Gaussian fit: 0.92

4.6 Discussion

In the previous chapter, I described the development of `quickPBSA` as data analysis framework for counting fluorescent emitters by photobleaching step analysis. After validation of the trace interpretation pipeline using semi-synthetic data with known ground-truth, I showed that `quickPBSA` provides robust emitter number estimates by measuring the number of fluorophores attached to DNA origami. These initial experiments suggested that the accessible counting range for `quickPBSA` was at least up to 35 emitters in experimental data and that the achieved speed-up in processing time required per trace was high enough to process large datasets ($>1 \times 10^3$ traces) within hours on a standard desktop computer.

To translate the findings regarding the performance of `quickPBSA` to counting of POI copy numbers in protein complexes within cells, I optimized the data acquisition conditions for cellular samples. This was done by comparing different fluorescent labels with respect to their photostability and tendency to undergo photoblinking between the fluorescent and transient dark states. The trace extraction approach was adapted by applying a filtering step to remove unsuited ROI from further analysis and by establishing an adaptive local background correction to address the largely increased sample complexity encountered in cells.

Testing two proteins for use as reference standard showed that NUP107, a component of the nuclear pore complex, is a suitable validation target for in situ measurements. Using a genome-edited U2OS cell line expressing NUP107-SNAP, I finally demonstrated that `quickPBSA` could be used for emitter counting on diffraction-limited structures inside eukaryotic cells. In the following, I will provide a discussion of technical aspects of relevance for the results described in this chapter.

In contrast to several other previously published approaches for PBSA, `quickPBSA` was validated with both, simulated and experimental data. Provide us with an objective picture of the performance that can be expected for target complexes with unknown numbers of protein copies. In addition, the benchmarking datasets generated in this study based on semi-synthetic traces with known ground-truth and from experimental data recorded on DNA origami, can also be used for benchmarking of alternative PBSA approaches. Here, a comprehensive comparison of different algorithms with respect to the achieved precision, bias and computational cost would help to make recommendations regarding the choice of PBSA approach for future biological studies. Benchmarking with semi-synthetic data and experimental validation recorded on DNA origami could also serve as base for an in-depth study of the dependence of `quickPBSA` performance on input data properties. Systematic variations of the temporal sampling of photobleaching traces, the number of data points and the SNR of traces can be performed in a controlled fashion using both types of data sources. This way, conditions which reduce the demand on data acquisition (and correspondingly the size of raw data) and potentially also conditions which improve the precision of emitter number estimates by `quickPBSA` could be identified.

No ground-truth information or independent validation with alternative emitter counting techniques was available for the in-situ emitter counting experiment, where I determined the number of NUP107-SNAP labeled with BG-SiR in individual NPCs inside eukaryotic cells (Section 4.5.4). However, in this case an independent validation of the results obtained with `quickPBSA` comes from labeling efficiency measurements described in chapter 3. Here, the number of NUP107-SNAP copies labeled with BG-SiR per NPC (10.9 ± 3.8, mean±SD) should correspond to the product of the known copy number of NUP107 per NPC (32x)[212] and the labeling efficiency. For labeling of SNAP$_f$-tag with BG-SiR at a concentration of 200 nM and a ligand incubation time of 2 h as was used for labeling of U2OS NUP107-SNAP, saturated labeling with an average labeling efficiency of $36 \pm 10\,\%$ (mean±SD) across different cell types was observed (Section 3.6.2, Table 3.1). Hence, the expected number of labeled NUP107-SNAP per NPC based on the known labeling efficiency of 10.9 ± 3.8 is in excellent agreement with the values reported by `quickPBSA`. In addition to this finding, which further supported that `quickPBSA` represents a promising tool for in-situ emitter counting by PBSA, further interesting observations were made in the course of this chapter.

In section 4.5.1, I showed that switching from fluorescent proteins as labels to self-labeling protein tags in combination with data acquisition in photostabilizing ROXS buffers drastically improved the quality of photobleaching traces. Based on the data presented above, it is likely that the full potential of PBSA for emitter counting has not yet been reached. Further optimization of labeling and data acquisition, e.g. by exchanging the fluorophores used here with fluorophores exhibiting a higher molecular brightness could lead to further improvements in trace quality and thereby help to extend the accessible counting range of PBSA. For example, in Section 3.6.3, I compared different ligands for SNAP$_f$-tag and HaloTag with respect to their achieved labeling efficiency and identified HTL-JF646 as a promising alternative ligand. In comparison to SiR, JF646 was reported to exhibit a 40 % increase in molecular brightness[183]. An interesting alternative approach for increasing the photostability and reducing photoblinking was developed in the form of so-called self-healing fluorophores[213]. These fluorophores are conjugated directly to redox reactive agents, which led to increased photostability and reduced photoblinking, while maintaining the molecular brightness of the fluorophore itself[214,215].

In this context, it is important to mention that the performed comparison of fluorophore photostabilities was performed at different excitation power densities for fluorophores excited at different wavelengths due to limitations in available excitation light sources. Subsequent normalization to an adjusted excitation power density enabled me to compare different fluorophores. However, it is known that the applied excitation power density can influence the rate of photobleaching in a non-linear fashion[216,217]. For this reason, a more detailed investigation of optimal excitation power densities, which maximize the photon output per fluorophore before irreversible photobleaching could help to maximize the trace quality for a chosen fluorophore and imaging condition.

In Section 4.5.2, I characterized the dodecameric *E. coli* enzyme complex GlnA and the nucleoporin NUP107 located in the nuclear pore complex of mammalian cells to determine if these proteins could serve as in-cell standard for benchmarking PBSA approaches. For the characterization of GlnA I first established a transgenic HeLa cell line stably transduced with GlnA-HaloTag by viral transduction. During subsequent experiments with this cell line, I observed that stable expression of GlnA-HaloTag was difficult to achieve although the chosen strategy for viral transduction allowed me to readily obtain stable HeLa gSEP and LynG as described in chapter 3 and even though selection of cells with the selection antibiotic puromycin was highly effective. One potential reason for this behavior might be the fact that GlnA is a metabolically active enzyme which catalyzes the condensation of glutamate into glutamine under Adenosine triphosphate (ATP) consumption. It might therefore be favorable for cells to suppress the expression of GlnA-HaloTag or formation of active complexes. In addition, the localization of GlnA in the cytosol led to a large variability of observed GlnA-HaloTag cluster numbers and intensities depending on the chosen condition during chemical fixation. Both of the observed characteristics, make GlnA-HaloTag a challenging in-cell reference standard with limited usability under selected conditions.

In contrast, NUP107-SNAP exhibited a largely different behavior in that it exhibited stable expression after genomic integration, low susceptibility towards stoichiometry changes during chemical fixation and, most importantly, a low complex intensity variability across complexes and across cells. However, emitter counting of NUP107-SNAP with `quickPBSA` revealed that the majority of ROIs used for extraction of photobleaching traces likely contained intensity contributions from two individual NPC. This was despite the use of the candidate ROI filtering module of `quickPBSA` and despite the adaptive background correction approach implemented in the trace extraction module (Section 4.5.3). Since the high density of NPC in the U2OS NUP107-SNAP cell line used here was already noticed during evaluation of the cell line, this finding is not entirely surprising and in line with previous reports that characterized the density and distribution of NPC on the nuclear envelope of eukaryotic cells[218–220]. A similar behavior was also reported when attempting to count the number of NUP133 copies in NPCs of U2OS cells with IBC and the fluorescent protein YFP as label[168]. One possible approach to reduce the crosstalk between adjacent nuclei would be to apply a more stringent candidate ROI filtering at the cost of reducing the number of traces obtained per cell.

Since in-cell standards are crucial for validation of quantitative fluorescence microscopy techniques, the observations made here motivate the development of alternative targets. Based on the observations made above, desirable features of candidates would be: a density of $\ll 1\,\mu\text{m}^{-2}$, absence of enzymatic activity in eukaryotic cells, absence of specific interaction partners in eukaryotic cells and the potential to be targeted to an organellar membrane or cytoskeletal filaments to facilitate efficient immobilization during fixation. A possible starting point for the evaluation of further reference standards is the prokaryotic 19.4 kDa protein ferritin (FtnA), which assembles into stable 24-mer complexes[221]. Ferritin is commonly used as standard in cryo-electron microscopy and was also suggested as standard in fluorescence microscopy[168]. Furthermore, targeting of the assembled ferritin oligomer to the membrane of cells by addition of a signaling peptide was demonstrated[222]. As alternative strategy an entirely new in-cell standard could be developed by building on recent advances in the design of artificial proteins[223,224].

Finally, the `quickPBSA` approach to PBSA can be directly compared with alternative approaches for emitter counting by fluorescence microscopy based on data acquire in the course of this study. In section 4.4, I used CoPS to characterize DNA origami, which were subsequently also used for emitter counting experiments with `quickPBSA` and the local intensity approach. These measurements can therefore serve as a direct comparison of the performance of all three approaches for measurements on in vitro samples. Here, CoPS, achieved a slightly lower variability (14.6 ± 5.0, mean±SD), than `quickPBSA` (14.2 ± 6.3, mean±SD). In contrast to `quickPBSA` and CoPS, emitter counting based on emission intensity along requires an additional intensity standard to relate intensities to absolute fluorophore numbers. To give an estimate about the precision that can be achieved with such an approach, the local intensity approach used above can be considered. For this approach, a variability comparable to that of `quickPBSA` (15.0 ± 6.4, mean±SD) was observed. However, this approach for intensity-based emitter number quantification does not rely on calibration with an external standard contained in a different sample and therefore likely was able to achieve a precision and bias better than what would typically be achieved with external standards. Finally, Thevathasan et al. recently performed emitter counting experiments with quantitative SMLM on nuclear pore complexes[92]. In these experiments, the copy number of NUP107 per NPC was determined by comparison with a reference standard in a second cell line and the cell-to-cell variability of the estimated number of NUP107 per NPC was found to be 8 %. For `quickPBSA` a similar degree of cell-to-cell variability of 4 % was observed for counting of NUP107-SNAP.

To conclude, our newly developed PBSA framework `quickPBSA` was shown to perform well in vitro and inside cells. As part of this work I have systematically optimized different parameters and validation standards to identify conditions that allowed us to perform successful PBSA measurements in cells. The analysis framework is openly available and together with the optimized labeling and imaging conditions described here can now serve as a validated tool for PBSA for future experiments and other researchers.

Part III

Conclusion and Future Perspectives

The characterization of macromolecular complexes with respect to their composition and function remains a key area of cell biological research. Ideally, information about the composition, stoichiometry and distribution of complexes within cells should be available to advance our understanding of complexes and their individual components. While fluorescence microscopy can in principle provide these information at the single-particle level, the need for fluorescent labeling and the need to convert fluorescent signals into absolute fluorophore numbers complicate quantitative assays aimed at elucidating complex stoichiometries. In this study, I presented two techniques that address these limitations of fluorescence microscopy and which, in combination, provide access to protein complex stoichiometry at the single-complex level and within cells.

In the first project presented in chapter 3, I described the implementation and validation of an improved version of the gSEP approach for measuring absolute labeling efficiencies by single-molecule colocalization analysis using the data analysis framework `DOLanalysis`. Based on this pipeline, I showed that the self-labeling protein tags SNAP$_f$-tag and HaloTag are limited in their ability to bind fluorescent ligands and that labeling efficiencies approaching 100 % could not be reached with either tag under the tested conditions. By systematically varying ligand concentration and reaction time, I found conditions where maximum labeling efficiency was achieved while minimizing non-specific background signal to negligible levels. This observation is very valuable, since labeling under these conditions will minimize the influence of slight changes in labeling conditions on achieved labeling efficiencies thereby increasing the robustness of labeling. Beyond this in-depth study of SNAP$_f$-tag and HaloTag, the gSEP approach now enables routine measurements of labeling efficiencies upon labeling of target proteins with self-labeling protein tags. This contributes to efforts to improve the quantitative rigor of fluorescence microscopy techniques for cell biology.

Even though I identified combinations of protein tag, fluorescent ligands and labeling conditions, which improved the achieved labeling efficiency to up to 65 %, the question what limits the maximal labeling efficiency for HaloTag and SNAP$_f$-tag remains. The fact that multiple reports by others support my findings suggests that more general mechanisms limit the ligand-binding efficiency of both tags. Identifying and potentially addressing the cause for limited labeling efficiencies for SNAP$_f$-tag and HaloTag is therefore a pressing issue for future studies given the widespread use of both tags and their otherwise very favorable properties for quantitative fluorescence microscopy. One possibility is that cellular factors might compete with the fluorescent ligand for the binding site in self-labeling protein tags.

Mass-spectrometry experiments could help to assess such a potential blocking of the catalytic center, by identifying cellular factors bound to individual amino acids in the catalytic center.

It was recently reported that fluorophores attached to SNAP$_f$-tag or HaloTag fused to different POI resulted in different signal intensities[207]. Based on this study, it remained unclear to which degree this effect can be attributed to changes in fluorophore brightness or labeling efficiency. In its current implementation gSEP reports on the labeling efficiency achieved in the cytosol of eukaryotic cells. However, the general approach of single-molecule colocalization analysis can in principle be extended towards other cellular organelles such as the endoplasmatic reticulum or mitochondria. For labeling efficiency measurements in other subcellular compartments a different image acquisition modality capable of single-molecule detection away from the coverslip might be required. Promising examples for potential microscopy techniques could be single-objective light sheet systems[225,226] or (spinning-disk) confocal microscopy[227].

The second project presented in chapter 4, described the development of the PBSA framework `quickPBSA` for emitter counting based on the identification of discrete intensity changes in intensity traces upon single-fluorophore photobleaching. The established approach was demonstrated to have an accessible counting range up to at least 35 emitters. In contrast to many previous approaches to PBSA, validation of `quickPBSA` was performed using simulated data as well as experimental data recorded from complexes with known stoichiometry and labeling efficiency. In comparison to a previous algorithm for advanced PBSA[123], the computational performance of `quickPBSA` was improved more than 100-fold without loss in accuracy, enabling future users to process large datasets on a standard desktop computer. A comparison with alternative emitter counting techniques further showed that `quickPBSA` achieves a precision comparable to that of emitter counting techniques based on photon antibunching and SMLM. As a calibration-free, robust and thoroughly validated emitter counting approach without the need for specialized instrumentation, `quickPBSA` puts routine measurements of protein complex stoichiometries in cell biological applications into perspective.

Overall, it is clear that in order to further improve emitter counting by PBSA, two complementary routes can be followed. First, it is likely possible to further improve the quality of input data by identifying brighter and more photostable fluorescent ligands for protein tags[228,229] or further optimization of buffer conditions during acquisition. A second route towards improved PBSA approaches lies in the possibility to develop new algorithms for trace interpretation and photobleaching step detection. In this respect, the semi-synthetic approach to quickly generate a large number of artificial photobleaching traces with high emitter numbers, realistic fluorophore behavior and known ground truth, opens exciting possibilities towards the development of novel approaches for trace interpretation based on machine learning. A recent report demonstrated that automated identification of photobleaching steps with convolutional neural networks is feasible[230]. While this approach was only able to discriminate between traces with up to four photobleaching steps, the availability of a large set of annotated ground truth data might enable to extend the counting range of such an approach.

On a broader level, the developments presented in the course of this work provide the basis for future studies in two general directions. First and foremost, the combination of calibrated labeling with self-labeling protein tags and fluorophore counting with `quickPBSA` provides a ready-to-use approach for measuring complex stoichiometries in biological samples with subcellular resolution and single-complex sensitivity. While quantitative fluorescence microscopy has become an increasingly used tool in cell biology during the past yearsrefs, many approaches are still difficult to implement and require specialized equipment and dedicated expertise. Both, the gSEP approach for measuring labeling efficiencies and the `quickPBSA` framework can be applied using equipment available to many non-expert laboratories through imaging facilities and thereby help to make quantitative fluorescence microscopy available to a broader user base.

While gSEP and `quickPBSA` were developed as two independent approaches, they readily benefit from each other due to the fact that both approaches were developed around self-labeling protein tags as excellent labels for single-molecule fluorescence microscopy. The labeling procedures and data acquisition conditions for SNAP$_\mathrm{f}$-tag or HaloTag identified in the course of this study provide a good starting point for establishing emitter counting assays based on gSEP and `quickPBSA` in cell biology. In this context, I showed that it is possible to determine the labeling efficiency for a given set of labeling conditions given a set of multi-color images from few ($\sim$10) cells expressing the gSEP construct. This opens up the possibility that gSEP cells used for calibration could be spiked into samples before labeling is performed to directly measure the labeling efficiency within a sample without the need for additional control samples.

While the developments presented in this study contribute to the growing toolbox available to fluorescence microscopists, further developments will be needed to transform fluorescence microscopy from a mainly qualitative technique into a truly quantitative approach, which readily reports on the quantity and distribution of individual molecular species.

One key aspect in that regard is the development of novel labeling strategies with further improved properties. While labeling kinetics and spectroscopic properties are readily available, the reachable labeling efficiency is often not well understood even for many labels commonly used in fluorescence microscopy today. In that respect, the gSEP approach represents a promising platform for comprehensive labeling efficiency studies since the `DOLanalysis` data analysis framework is compatible with a wide range of fluorescent labels and thereby ideally suited for comparing different labeling approaches. It will be interesting to see how alternative labels such as for example new protein tags[78,79,231,232] or peptide tags, which were recently reported to achieve relatively high labeling efficiencies[81,233] perform in systematic studies as were performed here for SNAP$_f$-tag and HaloTag. In this context, also the evaluation of labeling schemes based on bioorthogonal chemistry and genetic code expansion is a timely issue[234–236]. In the past years, a growing number of available chemistries for conjugation of small-molecule fluorophores to proteins via unnatural amino acids was reported[84,237–239]. However, for almost none of these approaches robust and comprehensive studies of labeling efficiencies that can be reached in live cells have been reported to date. Finally, the ongoing developments of new small-fluorophore ligands for SNAP$_f$-tag and HaloTag motivate further comprehensive studies of additional ligands. Many of these ligands were shown to exhibit favorable properties such as improved brightness or photostability and sometimes even fluorogenic properties upon binding to protein tags[229,240,241]. If high labeling efficiencies comparable to those achieved with SNAP$_f$-tag and the BG-SiR ligand can be reached, these labels would represent ideal candidates for emitter counting microscopy via PBSA and alternative emitter counting techniques such as counting by photon statistics[127].

Improved labeling efficiencies will also become increasingly important due to recent developments of new concepts for super-resolution fluorescence microscopy that are approaching molecular-scale resolution[242–246]. Here, the density of labels attached to the imaged structures will have to match the increase in resolution to facilitate sufficient sampling at high resolution[247–249]. For this reason, novel fluorescent labels with improved efficiency and concomitantly techniques for benchmarking labeling approaches will be critical.

As long as the ability of fluorescence microscopy to resolve the precise location of individual fluorophores in routine applications is limited to scales above the size of individual molecules, additional techniques to infer the number of fluorophores contained in dense structures from the fluorescence signal are required. This means that emitter counting techniques will likely remain important tools for cell biology in the years ahead. Still, developments towards integrating quantitative information obtained by emitter counting with additional information to improve the spatial resolution. Towards this goal, Bleaching and blinking-assisted localization microscopy (BALM) was developed as an alternative approach to SMLM[250]. While the current implementation of BALM provides images improved resolution, it might be possible to follow this approach to obtain high-resolution information in addition to emitter numbers. The fact that photobleaching in contrast to shelving of fluorophores in transient dark states is an irreversible process might even lead to advantages of BALM over other SMLM approaches for obtaining molecular numbers in addition to high-resolution images.

Future developments in fluorescence microscopy for biomedical research will largely depend on close collaboration across disciplines to come up with further improved labeling schemes, optical instrumentation and algorithms for data analysis. This study followed this approach by making use of fluorophore chemistry, single-molecule sensitive microscopy and automated data processing, to develop tools for quantitative studies of protein complexes, which will help to bring us one step further towards understanding the intricate processes that organize life at the nanoscale.

Part IV

Materials and Methods

Chapter 5

Materials

5.1 Reagents

5.1.1 Plasmids

All plasmids used in this study were previously used in peer-reviewed studies or were directly obtained from commercial sources. Table 5.1 provides an overview of all plasmids used and their respective sources.

Table 5.1 Plasmids used in this study.

Name	Source	Reference
pTOM20-HaloTag	Wombacher lab[1]	Hauke et al.[251]
pHaloTag-H2B	Wombacher lab	Hauke et al.[251]
pH2A-HaloTag	Wombacher lab	Hauke et al.[251]
pH2A-EGFP-HaloTag	Wombacher lab	
pTOMM20-mCherry-HaloTag	Wombacher lab	Werther et al.[252]
pmNeonGreen-TOMM20	Allele Biotech	Shaner et al.[206]
pGlnA-HaloTag	Heilemann lab[2]	Finan et al.[168]

---

[1] Max-Planck Institute for Medical Research, Heidelberg
[2] Goethe University Frankfurt

5.1.2 Cell lines

Table 5.2 provides an overview of transgenic cell lines used in this study. The procedures used to generate stable transgenic cell lines in the course of this study are described in Section 6.1.2. All cell lines expressing gSEP or LynG constructs as well as the HeLa GlnA-HaloTag cell line were cultivated under Puromycin selection ($1.5\,\mu g\,mL^{-1}$).

Wild-type HeLa cells were obtained from ATCC, wild-type Huh7.5 and H838 cell lines were obtained from the laboratory of Prof. Ursula Klingmüller (DKFZ Heidelberg) and U2OS wild-type cells were obtained from the laboratory of Dr. Jan Ellenberg (EMBL Heidelberg). COS-7 wild-type cells were obtained from the laboratory of Dr. Richard Wombacher (Max-Planck Institute for Medical Research, Heidelberg).

Table 5.2 Transgenic cell lines used in this study.

Parent line	Transgene	Source	Reference
Huh7.5	gSEP	Siegfried Hänselmann[1]	Dissertation[174]
Huh7.5	LynG	Siegfried Hänselmann	Dissertation[174]
HeLa	gSEP	Florian Salopiata[2]	Dissertation[253]
HeLa	LynG	Florian Salopiata	Dissertation[253]
H838	gSEP	Florian Salopiata	Dissertation[253]
H838	LynG	Florian Salopiata	Dissertation[253]
HeLa	GlnA-HaloTag	Florian Salopiata	
U2OS	NUP107-SNAP	Ellenberg lab[3]	Li et al.[210]

5.1.3 Protein tag ligands

Ligands for SNAP-tag or HaloTag were obtained from different commercial and non-commercial sources listed in Table 5.3. Janelia Fluor (JF) ligands for both tags were synthesized from N-Hydroxysuccinimide ester (NHS-Ester)-functionalized dyes (gift of Luke Lavis, Janelia Research Campus, US) and amine-functionalized BG or HTL linkers (Dominik Brox, Heidelberg University). First, 50 µL dye-NHS-Ester ($2\,mg\,mL^{-1}$) were mixed with 5 mg BG- or HTL-amine dissolved in 150 µL Dimethylformamide (DMF) and 20 µL N,N-Diisopropylethylamine (DIPEA) was added. Reaction mixtures were incubated

---

[1] Heidelberg University

[2] DKFZ Heidelberg

[3] EMBL Heidelberg

on a shaker for 3 h at room temperature. The formed conjugates were separated from educts via HPLC and the success of reaction was confirmed via mass spectrometry for all four conjugates. Purified conjugates were stored in DMF at −20 °C. Coupling, purification and characterization of HTL- or BG-functionalized JF549 and JF646 was performed by Felix Braun (Herten lab, Heidelberg University).

Table 5.3 Protein tag ligands used in this study.

Name	Alternative name	Source
HTL-TMR		Promega
HTL-SiR	SiR-Halo	Johnsson lab[1]
HTL-AF647	Halo-O2-AF647	Ries lab[2]
HTL-JF549		this study
HTL-JF646		this study
BG-TMR	TMR-Star	NEB
BG-SiR	SiR-SNAP	Spirochrome
BG-AF647		Ries lab
BG-JF549		this study
BG-JF646		this study

5.1.4 Fluorescently-labeled oligonucleotides

DNA origami labeled with AT647N or AT565 were obtained from Gattaquant DNA Nanotechnologies, dissolved in 0.5x TBE buffer (89 mM Tris base, 89 mM boric acid, 2 mM EDTA) supplemented with 11 mM $MgCl_2$ and stored at −20 °C until further use. tetraATTO647N or DNA oligonucleotides labeled with a single AT647N as described in Grussmayer and Herten[51] were obtained from biomers.net and stored in 1xPBS at −20 °C.

5.1.5 Chemicals & Consumables

All chemicals were obtained from Sigma-Aldrich unless indicated otherwise and stored according to manufacturer's recommendations. Reagents for cell culture were obtained from Gibco unless indicated otherwise.

---

[1] Max-Planck Institute for Medical Research, Heidelberg

[2] EMBL Heidelberg

5.2 Fluorescence microscopes

Widefield fluorescence microscopy was performed on two separate microscope setups described below. Unless noted otherwise, setup A was used for experiments.

5.2.1 Widefield setup A

Setup A was built around an inverted Nikon TI-Eclipse microscope stand equipped with a total internal reflection illumination unit, a motorized nosepiece, an autofocus system (PFS2, all Nikon) and a motorized XY-stage (Marzhäuser). A fiber-coupled multi laser engine (MLE-LFA, TOPTICA Photonics) equipped with 405 nm, 488 nm and 640 nm diode lasers and a 561 nm Diode pumped solid state (DPSS) was fed into the microscope stand through the TIRF illumination unit and used for sample excitation. A quadband dichroic filter (Di01-R405/488/561/635, Semrock) was used to direct excitation light towards a 100x 1.49 NA Apo TIRF oil immersion objective (Nikon) and to separate emitted light collected by the objective from excitation light. Remaining excitation light was further removed by directing the emitted light through a quadband notch filter (AHF Analysetechnik) placed between the dichroic filter and the tube lens (Nikon). Emitted light was then directed through a motorized filter wheel (FW102C, Thorlabs) and a 3-color image splitter (Optosplit III, Cairn) before being focused onto a back-illuminated EMCCD camera (iXon Ultra 897, Andor). The image splitter was equipped with exchangeable spectral beam splitters and bandpass filters to enable simultaneous multi-color imaging. The motorized filter wheel was equipped with 525(50) nm, 605(70) nm, and 690(70) nm bandpass filters (all AHF Analysetechnik) to enable automated multi-color imaging. The back-projected pixel size in the sample plane was 96 nm. Excitation lasers and camera were synchronized via Transistor–transistor logic (TTL) signals from the camera which were forwarded to the laser engine via a microcontroller board with analog in- and output channels (Arduino Uno, arduino.cc). All motorized components of the setup were controlled using the µManager software package[177]. To enable quantitative imaging, the camera's electron-multiplying (EM) gain as well as the output power of the 488 nm, 561 nm and 640 nm laser lines and their spatial illumination profiles were characterized (Fig. 5.1). The resulting maximum illumination power densities for each laser lines are given in Table 5.4.

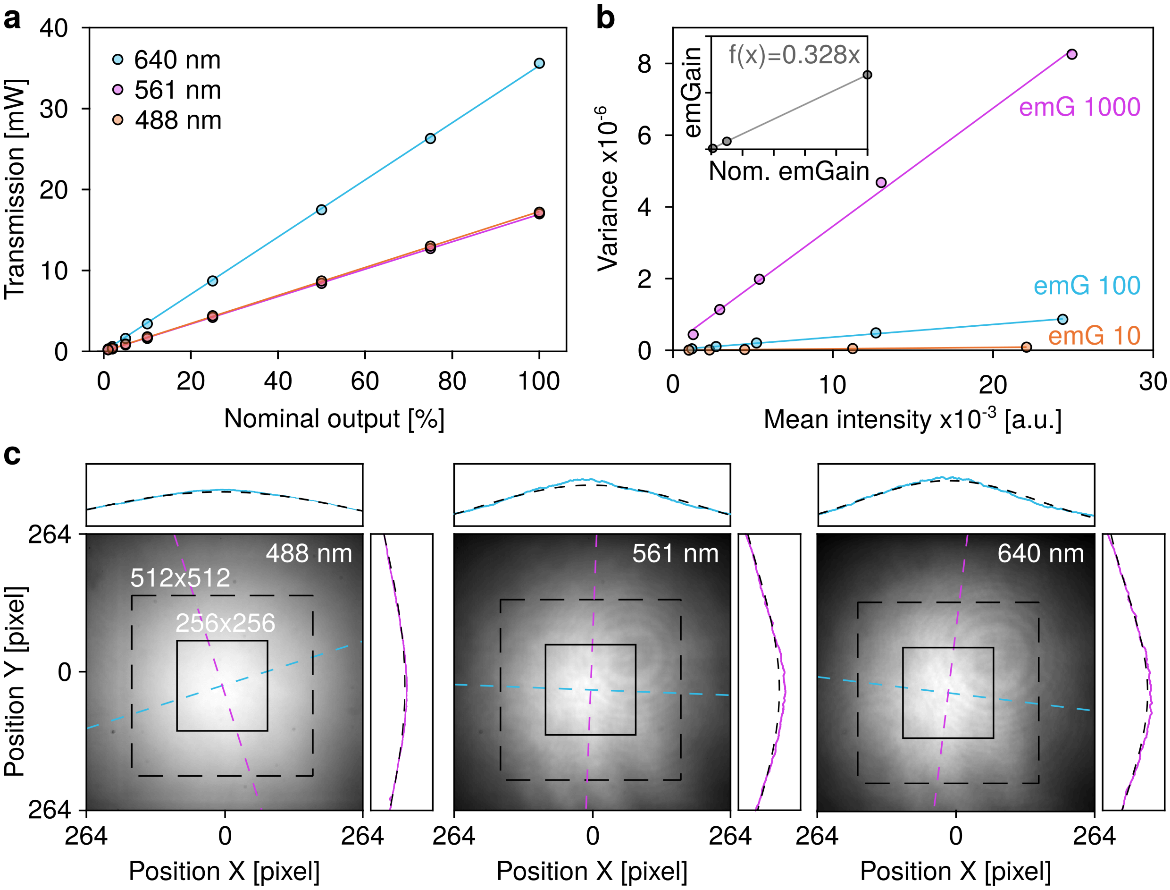

Figure 5.1 Characterization of illumination power, gain conversion and illumination profile for setup A. a, Transmitted excitation laser powers at different nominal laser output intensities. Linear fits to show linear dependence of measured and nominal laser intensities. **b,** The effective camera gain for the employed EMCCD at different nominal gain settings determined as described in Section 6.2.1. Inset: Linear fit of determined effective gain against nominal gain to show linear dependence of effective and nominal gain. R^2 values for all fits shown >0.99. **c,** Illumination profiles for different excitation lasers. Illumination profiles were fitted with a 2D Gaussian function. Axes and corresponding profiles along axes indicated in magenta and cyan. Dashed and solid lines indicate FOVs for full camera sensor (512x512 pixel) and central quarter of the camera sensor (256x256 pixel) using the 1.5x magnification tube lens, respectively.

Table 5.4 Illumination profile parameters of setup A. Full-width at half maximum (FWHM) values for axes as shown in Fig. 5.1. Fraction of illumination light within indicated regions in Fig. 5.1. Average illumination power densities were calculated using the transmitted laser power and the fraction of illumination light in the respective regions.

	488 nm	561 nm	640 nm
$\mathrm{FWHM_x}$ [µm]	97.4	74.0	57.5
$\mathrm{FWHM_y}$ [µm]	99.0	79.4	63.9
$\mathrm{Fraction_{256}}$ [%]	19.8	29.8	41.1
$\mathrm{Fraction_{512}}$ [%]	57.7	74.6	87.0
Illumination power $\mathrm{density_{256}}$ [$\mathrm{kW\,cm^{-2}}$]	0.56	0.83	2.39
Illumination power $\mathrm{density_{512}}$ [$\mathrm{kW\,cm^{-2}}$]	0.40	0.52	1.26

5.2.2 Widefield setup B

Setup B was built around an inverted Axiovert 200 microscope stand (Zeiss). A 647 nm Continuous-wave (CW) fiber laser (2RU-VFL-P-2000-647, MPB Communications) was used as excitation light source. The laser was first passed through a half wave plate (Thorlabs) and an Acousto-optic tunable filter (AOTF) (AOTFnC-Vis-TN, AA Opto-Electronic)and then expanded to 6.0 mm diameter ($1/e^2$). The Gaussian beam profile of the laser was converted into a top-hat profile using a refractive beam shaping device (πShaper 6 6 VIS, AdlOptica)[254] and further expanded to a diameter of 47 mm. The expanded and collimated beam was then focused onto the back focal plane of a 100x NA 1.49 oil immersion objective (Apo TIRF, Nikon) using a 500 mm achromatic doublet lens. Light emitted by the sample was collected through the same objective and separated from excitation light using a quad-band dichroic filter (R405/488/561/635, Semrock) which was also used to direct the excitation laser into the objective. Reflected excitation light was removed from the emission beam path by insertion of a quad-band notch filter (405/488/561/647, Semrock) and additional spectral filtering using a 700(50) nm bandpass filter (CHROMA). Finally, the remaining light was focused onto a back-illuminated scientific Complementary metal-oxide-semiconductor (sCMOS) camera (Prime 95B, Photometrics) using a 150 mm lens. The back-projected pixel size of 107.6 nm in the sample plane was determined using a micrometer ruler and the pixel size calibration plugin of µManager. The sample was placed on a motorized three-axis stage (MS2000) and kept in focus using an autofocus system (CRISP, both Applied Scientific Instrumentation) which used the back-reflection of a Light-emitting diode (LED) emitting at 850 nm to maintain the distance between coverslip and objective during image acquisition. The camera and the AOTF were synchronized using a microcontroller board (Arduino Mega, arduino.cc). All microscope components were controlled using the µManager software platform[177]. Microscope setup B was designed and implemented by Jonas Euchner under my supervision with help from Felix Braun.

5.2.3 Custom confocal scanning fluorescence microscope

CoPS measurements were performed on a custom-built confocal microscope. The setup was built around an inverted microscope stand (Axiovert 100, Zeiss), a ps-pulsed laser diode emitting at 640 nm (LDH P-C-640B, PicoQuant) operated at 20 MHz repetition rate and four single-photon sensitive Avalanche photodiodes (SPCM AQR-13, Perkin-Elmer). Linearly polarized light emitted by the laser diode was circularized using a quarter wave plate before being coupled into a single-mode polarization maintaining fiber (Schäfter Kirchhof). The excitation light was then directed towards a 100x NA 1.45 objective (Alpha Plan-Fluar, Zeiss) using a dichroic mirror (z532/640, CHROMA). Light emitted by the sample was collected by the same objective and passed through the dichroic filter. Scattered excitation light was removed using a quad-band notch filter (488/532/631-640 nm, AHF Analysetechnik). Remaining emitted light was then spatially filtered using a 100 µm pinhole placed in the focal plane between two achromatic doublet lenses (f =75 mm, Thorlabs) which in turn were place at $2f$ distance. All remaining light was split into four equal paths using three 50:50 beamsplitters and focused on the four APDs using achromatic doublet lenses (f =200 mm). 685(70) nm bandpass filters were place in front of each APD. Signals detected by the APDs were processed using a multichannel time-correlated single photon counting system (HydraHarp400) and the SymPhoTime 64 software platform (both PicoQuant). The position of objective and sample were controlled by a one-axis piezo scanner (P-721 PIFOC) and a two-axis piezo stage (7332CD, both Physik Instrumente). The two axis piezo stage was also used for stage scanning during image acquisition.

Chapter 6

Methods

6.1 Sample preparation

6.1.1 Cell culture

Unless stated otherwise, the protocols for culturing and cryo-preservation described below were used for all cell lines used in this study.

Adherent mammalian cells were cultured in complete growth medium consisting of phenol red-free Dulbecco's modified eagle medium (DMEM) supplemented with 10 % FBS (Australia origin), 20 mM and 1 mM sodium pyruvate (all Gibco/Thermo Fisher Scientific) at 37 °C, 5 % CO_2 and in humidified atmosphere. Either T-25 culture flasks (Sarstedt) or tissue culture dishes (TPP) were used for routine culturing. Upon reaching 80 % confluence or after three days, cells were subcultured by detaching cells from dishes and reseeding them into fresh dishes. Cells were detached by aspirating growth medium from cells, washing them once with 1xPBS and incubation with TrypLE™ Express (Gibco) for 3-5 min. TrypLE was inactivated by addition of two volumes complete growth medium and cells were pelleted by centrifugation for 5 min at 500x g. TrypLE and growth medium was aspirated from the pelleted cells and the pellet was resuspended in complete growth medium. Cells were then seeded at a density of $6\text{-}8 \times 10^3\,\mathrm{cm}^{-2}$. Stably transduced cells were additionally exposed to selection antibiotics indicated in Section 5.1.2.

For long-term storage of cell lines, cells were seeded into tissue culture dishes (TPP) and grown for at least 24 h and to ~80 % confluence in complete growth medium at 37 °C, 5 % CO_2 and in humidified atmosphere. Cells were then detached from dishes using TrypLE™

Express as described above and resuspended in complete growth medium containing 10 % DMSO. Cells were then gradually cooled to $-80\,°\mathrm{C}$ at a cooling rate of $-1\,°\mathrm{C\,min^{-1}}$ and transferred to the vapor phase of liquid nitrogen tanks for long-term storage.

Cryo-preserved cells were thawed by agitation of frozen vials in a water bath at 37 °C. Thawed cells were then plated in tissue culture dishes containing 10 volumes complete growth medium. The growth medium was replaced 24 h after plating.

6.1.2 Preparation of stable cell lines

Transgenic cell lines stably expressing gSEP or LynG were generated by retroviral transduction using the Phoenix ampho retroviral transduction[173] system and the plasmids pBABE-Lyn-HALO-linker-SNAPf (gSEP) and pBABE-Lyn-GFP (LynG) generated by Siegfried Haenselmann (Heidelberg University) together with Florian Salopiata (DKFZ Heidelberg). The pBabe plasmid backbone confers transgene expression under control of the 5’ long terminal repeat (LTR) of Moloney murine leukemia virus[255]. The Phoenix Ampho system employs a modified HEK293T cell line stably expressing the retroviral env and gag-pol proteins for helper-free production of replication-incompetent viral particles after transient transfection with plasmids carrying psi packaging sequences[173].

Plasmids were transfected into the Phoenix ampho virus packaging cell line by calcium phosphate precipitation for 6 h to produce replication incompetent virus particles later used for transduction. Virus was harvested 24 h after transfection by filtering through 0.22 µm syringe filters. Huh-7.5, H838 or HeLa cells were transduced by spin transduction at 340x g for 3 h. Selection of transduced cells with puromycin was started 24 h after transduction. Huh-7.5 cell lines gSEP and LynG were generated by Siegfried Haenselmann[174]. Transgenic H838 and HeLa gSEP and LynG cell lines were generated by Florian Salopiata[253]. Huh-7.5 and HeLa gSEP and LynG cell lines were sorted by FACS as described in Section 6.1.3 to obtain cell lines with expression levels suited for single-molecule imaging.

Full-length expression of gSEP and LynG constructs in the established cell-lines was verified by Western blot. Cells were lysed in RIPA buffer (0.05 M Tris pH 7.4, 150 mM NaCl, 2 mM EDTA, $2\,\mathrm{g\,L^{-1}}$ sodium deoxycholate, 1 mM Na_3VO_4, 5 mM NaF, 2 % NP40, $2\,\mathrm{\mu g\,mL^{-1}}$ AP, $200\,\mathrm{\mu g\,mL^{-1}}$ AEBSF) on a rotator for 30 min and by sonication on ice. Insoluble cell fragments were removed by centrifugation (20000x g, 10 min) and the supernatant was denatured in sample buffer ($64\,\mathrm{g\,L^{-1}}$ SDS, 40 mM Tris pH 7.4, 8 % glycerol, $12.3\,\mathrm{g\,L^{-1}}$ DTT, $0.16\,\mathrm{g\,L^{-1}}$ bromophenol blue, 20 % 2-mercaptoethanol) for 2 min at 95 °C. Samples were then separated on a 10 % sodium dodecyl sulfate polyacrylamide gel. For immunoblotting, denatured samples were transferred from SDS gels to polyvinylidene fluoride (PVDF) membranes while immersed in Laemmli buffer (14.4 $\mathrm{g\,L^{-1}}$ glycine, 3 $\mathrm{g\,L^{-1}}$ Tris, 1 $\mathrm{g\,L^{-1}}$ SDS) in a Hoefer TE 77 transfer unit (Amersham Biosciences) for 1 h at 260 mA. After transfer, membranes were blocked for 1 h in TBS-Tween buffer (150 mM NaCl, 10 mM Tris pH 7.4, 0.2 % Tween-20) supplemented with 5 % BSA and incubated with primary anti-GFP, anti-$\mathrm{His_6}$ (both Cell Signaling) or anti-actin (Sigma-Aldrich) in TBS-Tween supplemented with 1 % BSA overnight at 4 °C. The next day, membranes were washed in TBS-Tween, incubated with secondary antibodies conjugated with horseradish peroxidase (Dianova) for 1 h at room temperature and washed with TBS-Tween. Membranes were then covered with a 1:1 mixture of ECL buffers A (0.1 M Tris pH 8.5, 1.1 $\mathrm{mg\,L^{-1}}$ luminol, 0.185 $\mathrm{mg\,L^{-1}}$ p-coumaric acid, 1 % DMSO) and B (0.1 M Tris pH 8.5, 0.018 % H_2O_2) and incubated for 1 min. Chemiluminescence generated during oxidation of luminol by horseradish peroxidase was detected on a ImageQuant LAS4000 imaging system (GE Healthcare). Western blots were performed by Florian Salopiata.

A transgenic HeLa cell line stably expressing GlnA-HaloTag was established by excising GlnA-HaloTag from pGlnA-HaloTag and inserting it into the multiple cloning site of pBabe-M3-puro[253] to generate pBabe-GlnA-Halo. A stable HeLa cell line was then generated using the Phoenix ampho system as described above. All cloning and transductions for this cell line were performed by Florian Salopiata.

6.1.3 Flow cytometry

Analytical flow cytometry and FACS was performed with a FACSAria III cell sorter operated using FACSDiva software version 8.0.1 (both BD Biosciences). Before sorting, cells were detached from culture dishes as described above, resuspended in 1xPBS and additionally filtered through a cell strainer (Thermo Fisher Scientific) to remove cell aggregates. Gating for live cells was then performed based on front and side scatter signal resulting in approximately 40 % of events being discarded. Sorting was then performed based on fluorescence signal upon 488 or 633 nm excitation for EGFP- and SiR-labeled samples respectively. For analytical flow cytometry, typically 10×10^4 cells per condition were analyzed. Sorted cells were collected into tubes containing pre-warmed complete growth medium supplemented with penicillin (100 $\mathrm{U\,mL^{-1}}$) and streptomycin (100 $\mathrm{\mu g\,mL^{-1}}$). Cells were transferred to culture dishes as soon as possible and cultured in complete growth medium with penicillin and streptomycin for the next seven days. The flow cytometer was maintained and operated by the ZMBH flow cytometry & FACS core facility (Heidelberg University) and all experiments were performed under supervision of Dr. Monika Langlotz.

6.1.4 Labeling of cells expressing protein tag fusion proteins

Labeling of protein tags in living cells

Cells expressing SNAP_f-tag or HaloTag fusion proteins were labeled with varying ligands at variable concentrations and incubation times. Conditions used for specific experiments are stated in the description of individual experiments below. After labeling, cells were washed four times in full growth medium pre-warmed to 37 °C by carefully replacing all medium in a sample with fresh growth medium and incubating samples for 15 min (first two washes), 45 min (third wash) or 5 min (fourth wash) at 37 °C and 5 % CO_2. Samples were then chemically fixed as described below (Section 6.1.5).

Preparation of U2OS NUP107-SNAP samples

Genome edited U2OS cells expressing NUP107-SNAP were seeded into 8-well chambered coverslips (Nunc, Thermo Fisher Scientific) at a density of 8×10^3 cells per well in 250 µL complete growth medium and grown for 24 h. Cells were then labeled with 200 nM BG-SiR for 120 min at 37 °C. After labeling, cells were washed twice with complete growth medium, followed by three extended washes (2x30 min, 1x60 min) with complete growth medium at 37 °C. Samples were fixed in 3.7 % electron microscopy-grade PFA for 30 min as described in Section 6.1.5.

6.1.5 Chemical fixation of adherent cells

Before fixation, samples were washed once with pre-warmed 1xPBS. Samples were then fixed in 3.7 % PFA (v/v in 1xPBS) for 30 min followed by three washes (3x5 min) in 1xPBS. Fixative solutions were prepared by either preparing 3.7 % PFA solutions in 1xPBS by dissolving PFA powder in 1xPBS at 60 °C under addition of 1 M NaOH followed by adjusting the pH to 6.8 by addition of HCl after PFA had dissolved or by diluting a 16 % solution of electron microscopy grade PFA (electron microscopy grade, Science Services GmbH) with

5xPBS, and distilled H_2O directly before use. Fixative solutions were pre-warmed to 37° before addition onto samples.

All incubation steps during fixation were performed at room temperature and protected from light. Samples were either imaged directly after the last washing step or stored at 4 °C protected from light for up to 24 h before imaging.

6.1.6 Preparation of single-molecule surfaces

Samples for in vitro single-molecule fluorescence microscopy were prepared by specific immobilization of biotin-functionalized biomolecules onto chambered coverglasses. First, 8-well LabTek chambered coverglasses were cleaned by repeated incubation with 0.1 M hydrofluoric acid for 30-60 s followed by washing with millipore-filtered water (3x after each hydrofluoric acid incubation). Cleaned chambered coverglasses were incubated with a 20:1 mixture of BSA and biotinylated BSA (5 $\mathrm{mg\,mL^{-1}}$ total concentration in 1xPBS) for at least 30 min at room temperature (RT) or over night at 4 °C. Unbound BSA/$\mathrm{BSA_{biotin}}$ was removed by repeated washing with 1xPBS. BSA-coated surfaces were incubated with Streptavidin (100 $\mathrm{\mu g\,mL^{-1}}$ in 1xPBS) for 20 min at RT and washed three times with 1xPBS. Biotinylated biomolecules were added at increasing concentrations (typically 0.1-10 nM) and surface coverage was determined via single-molecule TIRF microscopy. After the desired surface coverage was reached, unbound probes were removed by repeated washing with 1xPBS. Final washes of single-molecule surfaces carrying DNA origami were washed in 1xPBS supplemented with 10 mM $MgCl_2$. Prepared surfaces were directly imaged or stored at 4 °C for up to 48 h.

6.1.7 Preparation of photostabilizing ROXS buffers

Photostabilizing buffers were prepared according to previously published procedures[129,201]. All buffers were prepared from a base solution (50 mM phosphate buffer, 13.5 mM KCl, 0.685 M NaCl, 10 mM $MgCl_2$, 12.5 % v/v glycerol) adjusted to pH 7.4. Before addition of ROXS reagents, oxygen was removed from the base solution by flowing Argon through it for at least 20 min. 1 mM paraquat dichloride and 1 mM ascorbic acid served as reducing/oxidizing agents and were freshly dissolved in 1xPBS before addition. For oxygen removal, 10 mM $NaSO_3$ (ROXS $NaSO_3$), 50 nM protocatechuate-3,4-deoxygenase (>3 $\mathrm{U\,mg^{-1}}$) and 2.5 mM of protocatechuic acid (ROXS PCD) or 0.66 MD-Glucose, 5 kU catalase, 40-80 U glucose oxidase and 1 mM Tris(2-carboxyethyl)phosphine (ROXS GodCat).

6.2 Microscopy

6.2.1 Characterization of microscopes

Transmitted power at varying wavelengths were measured with a handheld optical power meter (Thorlabs) placed in the optical beam path at the sample position with immersion oil and a type #1 cover glass placed on the objective. Illumination power densities were determined by measuring the transmitted power after the objective and normalization to the illuminated area. The illuminated area was determined by recording >10 images at different positions from a homogeneous dye solution after replacing the 1.5x magnification tube lens with a 1x tube lens to increase the back-projected pixel size. Images were averaged and fitted with a two-dimensional Gaussian.

The effective gain of the EMCCD camera used in widefield setup A was calibrated according to single point mean variance test following published procedures[1].

6.2.2 Counting by photon statistics

Fluorophore numbers per immobilized DNA origami were determined on DNA origami immobilized as described above **??**. 20-30 min before starting measurements, ROXS PCD buffer was added to samples and sample chambers were sealed using Parafilm while making sure that no air bubbles remained trapped between buffer and Parafilm. CoPS measurements were then performed on a custom confocal microscope described in Section 5.2.3. For each sample, overview scans with a pixel size of 80 nm were acquired at maximum scan speed and with an excitation intensity of 1 µW to select positions for pointCoPS measurements. Overview images were processed with a custom-written ImageJ macro making use of the `thunderSTORM` localization microscopy package[178] to determine the position of local maxima with sub-pixel precision. The coordinates of detected maxima were then transferred to SymPhoTime 64 using a Python script. Time-correlated single photon counting (TCSPC) data was acquired for 3 s at each coordinate at an excitation intensity of 10 µW for 3 s. A

---

[1] https://www.photometrics.com/learn/calculators

second overview scan after TCSPC data acquisition was performed to ensure that samples remained in focus during measurements.

6.2.3 Labeling efficiency measurements with gSEP

Variation of labeling conditions across incubation times, concentrations and cell lines

Experiments were performed with H838, HeLa or Huh7.5 cell lines stably expressing gSEP or LynG. Labeling was done live cells according to the protocol described above (section 6.1.4) with incubation times of 15, 30, 60 min, 3, 16 h and ligand concentrations of:

- BG-SiR: 0 nM, 0.5 nM, 5 nM, 25 nM, 50 nM, 250 nM, 500 nM, 1 250 nM
- HTL-TMR: 0 nM, 0.1 nM, 1 nM, 5 nM, 10 nM, 50 nM, 100 nM, 250 nM

Samples were fixed in 3.7 % PFA as described above (Section 6.1.5) and imaged in serum-free DMEM.

Evaluation of live-cell compatible fluorophore ligands

BG-TMR, HTL-SiR, BG/HTL-JF549 and BG/HTL-JF646 ligands were evaluated in Huh7.5 gSEP and LynG cells with samples being labeled before fixation and and fixation with 3.7 % PFA as described above (section 6.1.4). All ligands were tested at concentrations of 1 and 10 nM (HaloTag ligands) or 5 and 100 nM (SNAP_f-tag ligands). A constant incubation time of 30 min was used for all samples. Samples were fixed in 3.7 % PFA as described above and imaged in serum-free DMEM.

Comparison of pre- and post-fixation labeling

Experiments were performed with HeLa gSEP and LynG cell lines. HaloTag and SNAP_f-tag tag were labeled with varying concentrations of HTL-TMR or BG-SiR for 30 min in living cells or after fixation in 3.7 % PFA (w/v in 1xPBS) for 40 min. All washing steps were performed identically for both types of samples and as described in section 6.1.4. Images were imaged in serum-free DMEM.

AF647 ligands

Cells were seeded in 8-well LabTek chambered coverslips at a density of 6×10^3 cells per well in 250 µL complete growth medium 24 hours before fixation. Samples were fixed in a 3-step protocol described in Thevathasan et al.[92]. First, samples were pre-fixed in fixation buffer (2.4 % PFA (v/v in 1xPBS) for 30 s. Then, cells were permeabilized for 3 min in 0.4 % Triton X-100 (v/v in 1xPBS) and finally fixed for 30 min in 2.4 % PFA (v/v in 1xPBS). After fixation, samples were quenched in 100 mM NH_4Cl (w/v in 1xPBS) for 5 min and washed twice in 1xPBS for 5 min each. Before labeling, samples were incubated with Image-iT FX Signal Enhancer (ThermoFisher Scientific) diluted 1:2 in millipore-H_2O for 30 min.

Labeling with HTL-SiR, HTL-AF647 and BG-AF647 was performed in labeling buffer (0.5 % BSA, 1 µM DTT in 1xPBS) for 120 min at room-temperature, protected from light. After labeling, samples were washed three times with 1xPBS and three times with 1xPBS with 30 min incubations between buffer replacement. Samples were stored at 4 °C before imaging and imaged in PBS.

Data acquisition

Data was acquired on widefield setup A (Section 5.2.1) in TIRF mode. Unless stated otherwise, samples were imaged with the following settings: exposure time: 50 ms, emGain: 100, excitation intensities (measured at objective back focal plane): 10.6 mW (640 nm), 3.4 mW (561 nm), 6.9 mW (488 nm).

For each spectral channel, image stacks with 20 frames were acquired. Channels were acquired in reverse order of the excitation wavelength (from 640 nm to 488 nm) to avoid bleaching of fluorophores prior to imaging. Fluorophores in the 561 nm excitation channel were photobleached after image acquisition (2 min of constant 561 nm illumination at 17 mW) to reduce energy transfer by FRET from EGFP during acquisition of the 488 nm channel.

Data shown in Section 3.6.5 (AF647 ligands) was acquired at 200 ms exposure time, 130 emGain and with excitation intensities of 2.6 mW (488 nm) and 21.2 mW (640 nm).

6.2.4 Photobleaching step analysis with quickPBSA

Data for photobleaching step analysis was acquired in ROXS PCD buffer for all samples. DNA Origami data was acquired with 2.4 $\mathrm{kW\,cm^{-2}}$ average laser power density at 640 nm for R09, R20, R35 and with 0.84 kW cm^{-2} average laser power density at 561 nm for Y09 origami. NUP107 data was acquired with 1.2 kW cm^{-2} average laser power density at 640 nm. Exposure times were 50 or 200 ms for all measurements. Unless stated otherwise, all data was acquired on widefield setup A.

6.2.5 Fluorophore photostability characterization

The photostability of different fluorophores and the influence of ROXS buffers on the blinking and photostability of fluorophores was evaluated by recording time-lapse data from samples labeled with the corresponding fluorophore upon high intensity excitation. The stability of AT647N was evaluated using biotinylated DNA oligonucleotides labeled with AT647N immobilized as described in Section 6.1.6. TMR and SiR ligands for SNAP-tag and HaloTag were evaluated in fixed cells using cell lines expressing NUP107-SNAP or GlnA-HaloTag labeled as described in section 6.1.4.

The stability of EGFP, mCherry and mNeonGreen was evaluated in COS-7 cells transiently expressing H2A EGFP-HaloTag, TOMM20-mCherry-HaloTag or TOMM20-mNeonGreen (see Table 5.1) fixed 24 h after transfection. For each fluorophore, the stability in PBS pH 7.4 and the $NaSO_3$, ROXS PCD and ROXS GodCat buffer systems was tested with buffer compositions as described in Section 6.1.7. Prior to imaging, samples were washed once with PBS pH 7.4, sample chambers were filled with the respective buffer and sealed with Parafilm to minimize gas exchange during experiments.

Bleaching curves were acquired on widefield setup A for all buffer-fluorophore combinations described above. EGFP and mNeonGreen labeled structure were bleached at 0.58 kW cm^{-2} average power density, mCherry and TMR labeled samples were bleached at 0.84 kW cm^{-2} average power density and ATTO647N- and SiR-labeled samples were bleached at 2.42 kW cm^{-2} average power density. Image series were acquired with constant illumination until samples were fully bleached and the signal reached a plateau.

6.2.6 Fluorophore photoblinking characterization

For characterization of fluorophore blinking, HeLa gSEP cells were used as samples. For this, HeLa gSEP were grown in LabTek chambered coverslips and labeled with TMR or SiR ligands for HaloTag or SNAP$_f$-tag as described in section 6.1.4. Cells were incubated with 100 (HaloTag ligands) or 200 nM (SNAP$_f$-tag ligands) for 2 h. After labeling, cells were fixed as described in Section 6.1.5 with addition of 0.05 % glutaraldehyde to the fixation buffer. For recording of image time series, samples were prepared in either PBS or ROXS PCD buffer and individual cells were randomly selected for imaging. 2×10^3 frame image stacks were acquired with TIRF at an illumination power density of $0.05\,\mathrm{kW\,cm^{-2}}$ and with an exposure time of 25 ms on setup B.

6.3 Software and data analysis

6.3.1 Statistical analysis

All plotting and statistical tests were performed using MATLAB versions 2019a or 2020a. Unless stated otherwise, box plots span the interval from the 25th to the 75th percentile with the median indicated by a horizontal line within the box. Whiskers extend to 1.5x the interquartile range. Violin plots were generated using the `violinplot` function written by Bastian Bechtold[2] and are shown with a bandwidth corresponding to $\approx$10 % of the mean across all compared conditions. Box plots within violin plots reach from the 25th to the 75th percentile with the median shown as circle. If not indicated otherwise, values are reported as median± one standard deviation. R^2 values adjusted to the residual degrees of freedom ($R^2{}_{\mathrm{adj}}$) were computed using MATLAB[3].

6.3.2 Labeling efficiency analysis with DOLanalysis

All labeling efficiency measurements were analyzed using the `DOLanalysis` framework implemented in MATLAB and written by Felix Braun[4]. Raw images used as input for point detection were obtained by computing average projections of the first 10 frames of individual image stacks.

Cell segmentation

Segmentation to create binary masks for cells in individual images was performed with reference channel images as input using the ImageJ/Fiji script `processAverageIJ` dis-

---

[2] https://github.com/bastibe/Violinplot-Matlab

[3] https://de.mathworks.com/help/curvefit/evaluating-goodness-of-fit.html

[4] https://bitbucket.org/akherten/dol/src/master/

tributed as part of `DOLanalysis` as described in[174]. In summary, background was removed with the `Subtract Background` function (radius: 20 px), the image was bandpass filtered (settings: filter_large=512 filter_small=25 suppress=None tolerance=5 autoscale saturate), median filtered (settings: radius=30) and thresholded using the `AutoThreshold` function (algorithm: Triangle). Masks were exported as .tif images and later imported into MATLAB using `DOLanalysis`.

Manual classification

Density correction was found to be critical for obtaining reliable labeling efficiency estimates. In order for density correction to return reasonable results, it is important to make sure that input data meets quality criteria since simulations were used to determine the specific density correction factors (see section 6.3.2). To ensure that image sets included in a labeling efficiency measurement met the quality demands, image sets were curated and unsuited images were excluded from further analysis.

Based on the assumptions made for simulated data (constant point density across the entire area of a cell, proper segmentation of cell boundaries) used for determining density correction factors, two criteria were applied to decide whether individual images were discarded from further analysis. First, the segmented mask of cells should not include areas outside of cells because these surplus area would lead to underestimated detected point densities. Second, the point density across across the segmented area should be as homogeneous as possible to avoid over- or under-correction due to areas with low or high point density. In addition, images were excluded form further analysis if irregularities during image acquisition had occurred. This included image pairs where the basal membrane was not in focus, image pairs with uneven illumination across the field of view and cells with increased background levels.

Manual classification of images was performed with the `imageSetInspector` graphical user interface of `DOLanalysis` using the segmentation mask and reference channel images as evaluation data with images being displayed in random order.

Emitter localization

Emitter localization in averaged images from individual image channels, was performed using `u-track` as described previously[174] or with `thunderSTORM`. For emitter localization with `thunderSTORM`, automated analysis of full datasets was performed using the `Fiji/ImageJ` script `DOL_thunderSTORMlocalization_efficient.ijm` contained in `DOLanalysis` was used. The following default settings for emitter detection with `thunderSTORM` were applied. Image filtering: Wavelet filter, B-spline order: 3, B-spline scale: 2.0. Approximate localization of molecules: Local maximum method, peak intensity threshold: $2.0std(Wave.F1)$. Sub-pixel localization: Method: PSF Integrated Gaussian. Fitting radius: 3 pixel. Fitting method: Maximum likelihood. Initial sigma: 1.5, Multi-emitter fitting: enabled. Maximum number of molecules per fitting: 3.

For filtering of sigma values, sigma values from all localizations in a given spectral channel across all images in a dataset were grouped and binned with a bin width of 0.1. The bin with the highest number of localizations was used to center the filtering window. Unless stated otherwise, localizations with sigma values outside the interval ±50 % around the central bin were excluded from further analysis.

For all experiments shown in section 3.6, emitter localization was performed using `thunderSTORM` with post detection emitter filtering based on the distribution of localization sigma values.

Registration of multi-channel image sets

Registration of image channel pairs for individual multichannel images was performed by computing a projective transformation between detected emitters in two channels using the MATLAB function `fitgeotrans` as implemented in `DOLanalysis`. To improve robustness of registration, transformation parameters were only computed for images where more than 50 emitters were detected in both, the reference and target channels. Furthermore, transformations where the computed translation in x or y exceeded 3 pixels, the rotation exceeded 5° or the scaling factors exceeded 5 % were discarded. The parameters from successful registrations on a given set of images were averaged and the averaged transformation parameters were subsequently applied to all images in the set.

Distance threshold determination

Distance cutoffs for compute the specific degree of colocalization were computed according to a previously described procedure[174] implemented in `DOLanalysis`. A global distance cutoff value was determined for all imageset contained in an experiment.

Density correction

Density correction was performed as described in Equation 3.2.2 on 36. To obtain the required correction factors, CF_{slope} and CF_{offset}, I made use of simulations to generate images exhibiting diffraction-limited signals as described above (Section 3.4.2). I then performed emitter localization and filtering with settings identical to the settings used for processing experimental data (section 6.3.2), plotted the recall against the detected emitter density in the reference channel and modeled the obtained data with a linear fit (Fig. 6.1).

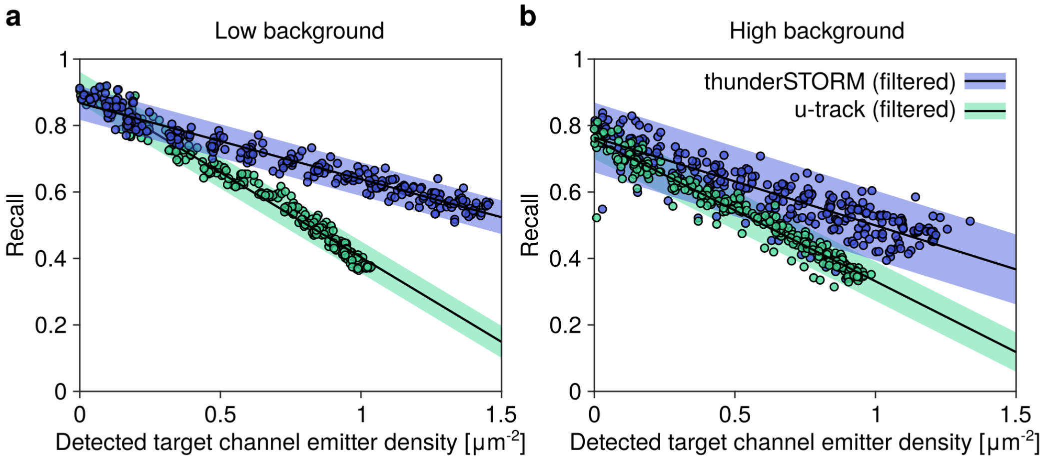

Figure 6.1 Density correction factors for raw DOL values determined from simulations with different background levels. Emitter images at defined densities ranging from 0 to 2.6 µm^{-2} were simulated and combined with low (a) or high (b) background images as described in Section 3.4.2. 20 generated images per condition were processed with the optimized point detection (`thunderSTORM` filtered) or previous (`u-track` filtered) point detection approaches and results were compared against ground truth emitter positions to compute recall (circles). Linear fits (black lines) were performed to obtain the density correction factor C_{dens} and offset value. Shaded regions represent 95 % confidence interval from fit. Simulations were performed by Felix Braun.

The obtained correction factors for emitter localization with `u-track` and `thunderSTORM` in the low and high background scenarios are given in Table 6.1.

Apparent rate constant fits

Median DOL values across varying labeling conditions were fitted to Equation 3.5 by non-linear least squares regression with the Trust-Region algorithm as provided by the MATLAB function `fit`.

Table 6.1 Correction factors obtained from analysis of simulated data and used for density correction after emitter localization with indicated localization frameworks. `thunderSTORM`. See section 6.3.2 for specific emitter localization settings.

Localization framework	BG	CF_{slope}	CF_{offset}
`thunderSTORM`	low	-0.2131	0.8681
`thunderSTORM` (filtered)	low	-0.2292	0.8677
`thunderSTORM`	high	-0.2383	0.7943
`thunderSTORM` (filtered)	high	-0.2647	0.7641
`u-track`	low	-0.4782	1.0119
`u-track` (filtered)	low	-0.51	0.9138
`u-track`	high	-0.4215	0.9470
`u-track` (filtered)	high	-0.4265	0.7579

6.3.3 Simulations for labeling efficiency measurements

Single emitter signal randomly distributed across images without background were simulated with testSTORM[179]. For this, emitter signal were generated using a Gaussian PSF model, the 'vesicle' pattern and acquisition parameters matching those of experimental data acquired in the scope of this thesis (section 6.2.3). No sample drift was included in simulated data. The dye model 'Alexa Fluor 647' with the following parameter values was chosen to adapt emitter properties to those observed in experimental data: on time: 0.025, off time: 0.01, bleaching constant: 0.2, emitted photon/s: 350, number of labels per epitope: 1.

6.3.4 CoPS data analysis

TCSPC data was analyzed by first computing coincident photon histograms for each sampled position from the initial 1×10^7 laser cycles of the corresponding TCSPC data stream. Histograms were then fitted to an analytical model of coincident photon event probabilities using the `scipy` implementation of the Levenberg-Marquardt algorithm contained in the `pycops` package[1] developed by Johan Hummert (Herten lab, University of Birmingham).

---

[1] https://bitbucket.org/akherten/pycops/

6.3.5 Data analysis with quickPBSA

The first five images from the measured image sequence were averaged and used to locate diffraction-limited signals in images. The localization was performed with `Fiji`[256] version 1.52p using the plugin `thunderSTORM`[178]. Trace extraction was done with the trace extraction module `quickPBSA`. In short, the average signal from circular regions around the localizations was extracted with typical diameters of 950 nm for in vitro samples and 150 nm for NUP107 samples. For background correction, the average signal from ring-shaped regions was subtracted (inner diameter 1.7 µm for DNA origami, 0.6 µm for NUP107 samples, outer diameter 2.0 µm for origami, 0.9 µm for NUP107). Regions around neighboring localizations were excluded from the background region. Additionally, ROIs with nearest neighbors at a distance below 950 nm for DNA origami samples and 475 nm for NUP107 samples were excluded. Photobleaching step analysis was performed using the `quickPBSA` package. Typically, the *threshold* parameter for preliminary step detection was set at 0.03 and *maxiter* at 200. Other analysis parameters were kept at their default values, except for the *mult_threshold* parameter in step refinement, which was typically set to 1.5 to decrease runtime. The analysis according to Teskouras et al.[123] was performed with available `python` code[5]. Data analysis was performed on a desktop computer with an 8-core CPU (Intel Core i7-3770, 3.4 GHz) and 12 GB DDR3 memory. Semi-synthetic datasets were generated by manual annotation of traces obtained from tetraATTO647N DNA oligonucleotides measured in $NaSO_3$ ROXS buffer. For visual inspection of `quickPBSA` trace models, models and intensity traces were overlayed by applying a linear scaling factor $y = ax$ to the quickPBSA model trace.

Distributions of fluorophore numbers measured on NUP107-SNAP samples and shown in Fig. 4.31 were modeled with the following equation:

$$f(x) = frac_{single} e^{-\frac{(x0-\mu)^2}{2\sigma^2}} + frac_{double} e^{-\frac{(x-2\mu)^2}{4\sigma^2}} \tag{6.1}$$

Where $frac_{single}$ and $frac_{double}$ are the fractions of ROIs with one or two NPC, μ is the average number of fluorophores per NPC and σ is the standard deviation. Fitting was performed with MATLAB by nonlinear least squares regression with the Trust-Region algorithm as provided by the MATLAB function `fit`.

---

[5] https://github.com/lavrys/Photobleach

6.3.6 Analysis of fluorophore photostability

Data acquired as described in **??** was background corrected by subtracting a constant offset from acquired image series to account for camera offset and excitation light bleed-through. Offsets were manually determined for each sample and were found to be well reproducible within one condition, but variable across conditions. Bleach curves were then extracted from image series by extracting the frame-wise average intensity within a masked region. Masks were obtained from a Gaussian-filtered average projection of the first 10 images in the series and local thresholding following the Bernsen method. Mask segmentation and intensity extraction was performed using custom-written code in `Fiji/ImageJ`. Bleach curves were normalized to the maximum intensity in the respective trace and the raw half bleach time ($\mathrm{t_{t1/2,raw}}$) defined as the time at which the intensity traces had decayed to $<50\,\%$ of the maximum intensity was extracted using custom-written MATLAB code. To facilitate comparison between fluorophores excited at different wavelengths, $\mathrm{t_{t1/2,raw}}$ were normalized against the applied illumination power density (IPD) to obtain the excitation power-corrected $t_{t1/2} = t_{1/2,raw}/IPD$ at an illumination power density of $1.0\,\mathrm{kW\,cm^{-2}}$.

6.3.7 Characterization of fluorophore photoblinking

Data acquired as described in Section 6.2.6 was used for analysis. First, images were correlated following the approach developed by Sehayek et al.[208] using MATLAB code provided by the authors[6]. ACFs firstplural were computed using the first 1×10^3 frames of each image stack and a maximum lag time of 6×10^2 frames. The initial decay (lag times 2-125 frames) of obtained ACFs firstplural containing information about photoblinking were fitted to a 3 state model (equalBleach) developed by Sehayek et al.[208].

---

[6] https://github.com/ssehayek/blink-project.git

Part V

Appendices

References

[1] Sebastian E. Ahnert et al. “Principles of Assembly Reveal a Periodic Table of Protein Complexes”. In: *Science* 350.6266 (2015), aaa2245. DOI: 10.1126/science.aaa2245.

[2] Kevin Drew et al. “Integration of over 9,000 Mass Spectrometry Experiments Builds a Global Map of Human Protein Complexes”. In: *Molecular Systems Biology* 13.6 (2017), p. 932. DOI: 10.15252/msb.20167490.

[3] Pierre C. Havugimana et al. “A Census of Human Soluble Protein Complexes”. In: *Cell* 150.5 (2012), pp. 1068–1081. DOI: 10.1016/j.cell.2012.08.011.

[4] Joseph A. Marsh and Sarah A. Teichmann. “Structure, Dynamics, Assembly, and Evolution of Protein Complexes”. In: *Annual Review of Biochemistry* 84.1 (2015), pp. 551–575. DOI: 10.1146/annurev-biochem-060614-034142.

[5] Martin Beck and Ed Hurt. “The Nuclear Pore Complex: Understanding Its Function through Structural Insight”. In: *Nature Reviews Molecular Cell Biology* 18.2 (2017), pp. 73–89. DOI: 10.1038/nrm.2016.147.

[6] Daniel H. Lin and André Hoelz. “The Structure of the Nuclear Pore Complex (An Update)”. In: *Annual review of biochemistry* 88 (2019), pp. 725–783. DOI: 10.1146/annurev-biochem-062917-011901.

[7] Vilma Jimenez Sabinina et al. “3D Super-Resolution Fluorescence Microscopy Maps the Variable Molecular Architecture of the Nuclear Pore Complex”. In: *Molecular Biology of the Cell* (2021), mbc.E20-11–0728. DOI: 10.1091/mbc.E20-11-0728.

[8] Anthony P. Schuller et al. “The Cellular Environment Shapes the Nuclear Pore Complex Architecture”. In: *Nature* 598.7882 (7882 2021), pp. 667–671. DOI: 10.1038/s41586-021-03985-3.

[9] Ioanna Bethani et al. “Spatial Organization of Transmembrane Receptor Signalling”. In: *The EMBO journal* 29.16 (2010), pp. 2677–2688. DOI: 10.1038/emboj.2010.175.

[10] Hao Wu. “Higher-Order Assemblies in a New Paradigm of Signal Transduction”. In: *Cell* 153.2 (2013), pp. 287–292. DOI: 10.1016/j.cell.2013.03.013.

[11] Jesse Goyette and Katharina Gaus. “Mechanisms of Protein Nanoscale Clustering”. In: *Current Opinion in Cell Biology* 44 (2017), pp. 86–92. DOI: 10.1016/j.ceb.2016.09.004.

[12] Michael L. Dustin and Jay T. Groves. “Receptor Signaling Clusters in the Immune Synapse”. In: *Annual Review of Biophysics* 41 (2012), pp. 543–556. DOI: 10.1146/annurev-biophys-042910-155238.

[13] Wolfgang W. A. Schamel and Balbino Alarcón. “Organization of the Resting TCR in Nanoscale Oligomers”. In: *Immunological Reviews* 251.1 (2013), pp. 13–20. DOI: 10.1111/imr.12019.

[14] Pieta K. Mattila et al. “The Actin and Tetraspanin Networks Organize Receptor Nanoclusters to Regulate B Cell Receptor-Mediated Signaling”. In: *Immunity* 38.3 (2013), pp. 461–474. DOI: 10.1016/j.immuni.2012.11.019.

[15] Jesse Goyette et al. “How Does T Cell Receptor Clustering Impact on Signal Transduction?” In: *J Cell Sci* 132.4 (2019), jcs226423. DOI: 10.1242/jcs.226423.

[16] Erdinc Sezgin et al. “The Mystery of Membrane Organization: Composition, Regulation and Roles of Lipid Rafts”. In: *Nature Reviews Molecular Cell Biology* 18.6 (6 2017), pp. 361–374. DOI: 10.1038/nrm.2017.16.

[17] Lindsay B. Case, Jonathon A. Ditlev, and Michael K. Rosen. “Regulation of Transmembrane Signaling by Phase Separation”. In: *Annual Review of Biophysics* 48 (2019), pp. 465–494. DOI: 10.1146/annurev-biophys-052118-115534.

[18] Kevin M. Dean and Amy E. Palmer. “Advances in Fluorescence Labeling Strategies for Dynamic Cellular Imaging”. In: *Nature Chemical Biology* 10.7 (2014), pp. 512–523. DOI: 10.1038/nchembio.1556.

[19] K. S. Grussmayer, K. Yserentant, and D.-P. Herten. “Photons in - Numbers out: Perspectives in Quantitative Fluorescence Microscopy for in Situ Protein Counting”. In: *Methods and Applications in Fluorescence* 7.1 (2019), p. 012003. DOI: 10.1088/2050-6120/aaf2eb.

[20] Joseph R. Lakowicz. *Principles of Fluorescence Spectroscopy*. Auflage: 3rd ed. 2006. Corr. 5th printing 2010. New York: Springer, 2010. 954 pp. ISBN: 978-0-387-31278-1.

[21] Bernard Valeur and Mário Nuno Berberan-Santos. *Molecular Fluorescence: Principles and Applications*. 2. edition. Weinheim, Germany : Chichester, England: Wiley-VCH, 2012. 592 pp. ISBN: 978-3-527-32837-6.

[22] George Gabriel Stokes. “XXX. On the Change of Refrangibility of Light”. In: *Philosophical Transactions of the Royal Society of London* 142 (1852), pp. 463–562. DOI: 10.1098/rstl.1852.0022.

[23] Michael Kasha. “Characterization of Electronic Transitions in Complex Molecules”. In: *Discussions of the Faraday Society* 9.0 (1950), pp. 14–19. DOI: 10.1039/DF9500900014.

[24] G. W. Byers, S. Gross, and P. M. Henrichs. “Direct and Sensitized Photooxidation of Cyanine Dyes”. In: *Photochemistry and Photobiology* 23.1 (1976), pp. 37–43. DOI: 10.1111/j.1751-1097.1976.tb06768.x.

[25] Marcia Levitus and Suman Ranjit. “Cyanine Dyes in Biophysical Research: The Photophysics of Polymethine Fluorescent Dyes in Biomolecular Environments”. In: *Quarterly Reviews of Biophysics* 44.1 (2011), pp. 123–151. DOI: 10.1017/S0033583510000247.

[26] Edward A. Lemke et al. “Microfluidic Device for Single-Molecule Experiments with Enhanced Photostability”. In: *Journal of the American Chemical Society* 131.38 (2009), pp. 13610–13612. DOI: 10.1021/ja9027023.

[27] Colin Echeverría Aitken, R. Andrew Marshall, and Joseph D. Puglisi. “An Oxygen Scavenging System for Improvement of Dye Stability in Single-Molecule Fluorescence Experiments”. In: *Biophysical Journal* 94.5 (2008), pp. 1826–1835. DOI: 10.1529/biophysj.107.117689.

[28] Marko Swoboda et al. “Enzymatic Oxygen Scavenging for Photostability without pH Drop in Single-Molecule Experiments”. In: *ACS Nano* 6.7 (2012), pp. 6364–6369. DOI: 10.1021/nn301895c.

[29] Qinsi Zheng et al. “The Contribution of Reactive Oxygen Species to the Photobleaching of Organic Fluorophores”. In: *Photochemistry and Photobiology* 90.2 (2014), pp. 448–454. DOI: 10.1111/php.12204.

[30] Ivan Rasnik, Sean A. McKinney, and Taekjip Ha. “Nonblinking and Long-Lasting Single-Molecule Fluorescence Imaging”. In: *Nature Methods* 3.11 (2006), pp. 891–893. DOI: 10.1038/nmeth934.

[31] Jerker Widengren et al. “Strategies to Improve Photostabilities in Ultrasensitive Fluorescence Spectroscopy”. In: *The Journal of Physical Chemistry A* 111.3 (2007), pp. 429–440. DOI: 10.1021/jp0646325.

[32] Jan Vogelsang et al. “A Reducing and Oxidizing System Minimizes Photobleaching and Blinking of Fluorescent Dyes”. In: *Angewandte Chemie International Edition* 47.29 (2008), pp. 5465–5469. DOI: 10.1002/anie.200801518.

[33] Robert Hooke. *Micrographia: Or Some Physiological Descriptions of Minute Bodies Made by Magnifying Glasses, with Observations and Inquiries Thereupon*. London: Royal Society, 1665. 390 pp. ISBN: 978-0-486-49564-4.

[34] Adam J. M. Wollman et al. “From Animaculum to Single Molecules: 300 Years of the Light Microscope”. In: *Open Biology* 5.4 (), p. 150019. DOI: 10.1098/rsob.150019.

[35] Peter Kapusta, Michael Wahl, and Rainer Erdmann. *Advanced Photon Counting: Applications, Methods, Instrumentation*. Softcover reprint of the original 1st ed. 2015 edition. Cham: Springer, 2016. 384 pp. ISBN: 978-3-319-35429-3.

[36] E. Abbe. “Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung”. In: *Archiv für mikroskopische Anatomie* 9.1 (1873), pp. 413–418. DOI: 10.1007/BF02956173.

[37] Bo Huang, Mark Bates, and Xiaowei Zhuang. “Super Resolution Fluorescence Microscopy”. In: *Annual review of biochemistry* 78 (2009), pp. 993–1016. DOI: 10.1146/annurev.biochem.77.061906.092014.

[38] T. Schmidt et al. “Imaging of Single Molecule Diffusion”. In: *Proceedings of the National Academy of Sciences of the United States of America* 93.7 (1996), pp. 2926–2929.

[39] Russell E. Thompson, Daniel R. Larson, and Watt W. Webb. “Precise Nanometer Localization Analysis for Individual Fluorescent Probes”. In: *Biophysical Journal* 82.5 (2002), pp. 2775–2783. DOI: 10.1016/S0006-3495(02)75618-X.

[40] Bo Huang et al. “Three-Dimensional Super-Resolution Imaging by Stochastic Optical Reconstruction Microscopy”. In: *Science* 319.5864 (2008), pp. 810–813. DOI: 10.1126/science.1153529.

[41] Hao Shen et al. “Single Particle Tracking: From Theory to Biophysical Applications”. In: *Chemical Reviews* 117.11 (2017), pp. 7331–7376. DOI: 10.1021/acs.chemrev.6b00815.

[42] Michael J. Rust, Mark Bates, and Xiaowei Zhuang. “Sub-Diffraction-Limit Imaging by Stochastic Optical Reconstruction Microscopy (STORM)”. In: *Nature Methods* 3.10 (2006), pp. 793–795. DOI: 10.1038/nmeth929.

[43] Mickaël Lelek et al. “Single-Molecule Localization Microscopy”. In: *Nature Reviews Methods Primers* 1.1 (1 2021), pp. 1–27. DOI: 10.1038/s43586-021-00038-x.

[44] Marta Fernández-Suárez and Alice Y. Ting. “Fluorescent Probes for Super-Resolution Imaging in Living Cells”. In: *Nature Reviews Molecular Cell Biology* 9.12 (2008), pp. 929–943. DOI: 10.1038/nrm2531.

[45] Karin Nienhaus and G. Ulrich Nienhaus. “Fluorescent Proteins for Live-Cell Imaging with Super-Resolution”. In: *Chemical Society Reviews* 43.4 (2014), pp. 1088–1106. DOI: 10.1039/c3cs60171d.

[46] E Wulf et al. “Fluorescent Phallotoxin, a Tool for the Visualization of Cellular Actin.” In: *Proceedings of the National Academy of Sciences of the United States of America* 76.9 (1979), pp. 4498–4502. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC411604/ (visited on 01/02/2022).

[47] Joëlle Dubois et al. “Fluorescent and Biotinylated Analogues of Docetaxel: Synthesis and Biological Evaluation”. In: *Bioorganic & Medicinal Chemistry* 3.10 (1995), pp. 1357–1368. DOI: 10.1016/0968-0896(95)00115-W.

[48] Charlotte Stadler et al. “Immunofluorescence and Fluorescent-Protein Tagging Show High Correlation for Protein Localization in Mammalian Cells”. In: *Nature Methods* 10.4 (2013), pp. 315–323. DOI: 10.1038/nmeth.2377.

[49] Peter J. Thul et al. “A Subcellular Map of the Human Proteome”. In: *Science* 356.6340 (2017), eaal3321. DOI: 10.1126/science.aal3321.

[50] R. P. Haugland. “Coupling of Monoclonal Antibodies with Fluorophores”. In: *Methods in Molecular Biology* 45 (1995), pp. 205–221. DOI: 10.1385/0-89603-308-2:205.

[51] Kristin S. Grußmayer, Anton Kurz, and Dirk-Peter Herten. “Single-Molecule Studies on the Label Number Distribution of Fluorescent Markers”. In: *ChemPhysChem* 15.4 (2014), pp. 734–742. DOI: 10.1002/cphc.201300840.

[52] A. Coulter and R. Harris. “Simplified Preparation of Rabbit Fab Fragments”. In: *Journal of Immunological Methods* 59.2 (1983), pp. 199–203. DOI: 10.1016/0022-1759(83)90031-5.

[53] Ágnes Szabó et al. “The Effect of Fluorophore Conjugation on Antibody Affinity and the Photophysical Properties of Dyes”. In: *Biophysical Journal* 114.3 (2018), pp. 688–700. DOI: 10.1016/j.bpj.2017.12.011.

[54] Ulrich Rothbauer et al. “Targeting and Tracing Antigens in Live Cells with Fluorescent Nanobodies”. In: *Nature Methods* 3.11 (2006), pp. 887–889. DOI: 10.1038/nmeth953.

[55] Valentin Fabricius et al. “Rapid and Efficient C-terminal Labeling of Nanobodies for DNA-PAINT”. In: *Journal of Physics D: Applied Physics* 51.47 (2018), p. 474005. DOI: 10.1088/1361-6463/aae0e2.

[56] Peter C. Fridy et al. “A Robust Pipeline for Rapid Production of Versatile Nanobody Repertoires”. In: *Nature Methods* 11.12 (2014), pp. 1253–1260. DOI: 10.1038/nmeth.3170.

[57] Felipe Opazo et al. “Aptamers as Potential Tools for Super-Resolution Microscopy”. In: *Nature Methods* 9.10 (2012), pp. 938–939. DOI: 10.1038/nmeth.2179.

[58] Christian Tiede et al. “Affimer Proteins Are Versatile and Renewable Affinity Reagents”. In: *eLife* 6 (2017). Ed. by Franck Perez, e24903. DOI: 10.7554/eLife.24903.

[59] Mariia Darmostuk et al. “Current Approaches in SELEX: An Update to Aptamer Selection Technology”. In: *Biotechnology Advances*. BioTech 2014 and 6th Czech-Swiss Biotechnology Symposium 33 (6, Part 2 2015), pp. 1141–1161. DOI: 10.1016/j.biotechadv.2015.02.008.

[60] Anna Ah-San Tang et al. “Isolation of Isoform-Specific Binding Proteins (Affimers) by Phage Display Using Negative Selection”. In: *Science Signaling* (2017). DOI: 10.1126/scisignal.aan0868.

[61] Sebastian Strauss et al. “Modified Aptamers Enable Quantitative Sub-10-Nm Cellular DNA-PAINT Imaging”. In: *Nature Methods* (2018), p. 1. DOI: 10.1038/s41592-018-0105-0.

[62] Thomas Schlichthaerle et al. “Site-Specific Labeling of Affimers for DNA-PAINT Microscopy”. In: *Angewandte Chemie International Edition* 0 (ja 2018). DOI: 10.1002/anie.201804020.

[63] Jennifer A. Doudna and Emmanuelle Charpentier. “The New Frontier of Genome Engineering with CRISPR-Cas9”. In: *Science* 346.6213 (2014), p. 1258096. DOI: 10.1126/science.1258096.

[64] Talley J. Lambert. “FPbase: A Community-Editable Fluorescent Protein Database”. In: *Nature Methods* 16.4 (2019), p. 277. DOI: 10.1038/s41592-019-0352-8.

[65] Richard Wombacher and Virginia W. Cornish. “Chemical Tags: Applications in Live Cell Fluorescence Imaging”. In: *Journal of Biophotonics* 4.6 (2011), pp. 391–402. DOI: 10.1002/jbio.201100018.

[66] Anselm Fabian Lowell Schneider and Christian Peter Richard Hackenberger. “Fluorescent Labelling in Living Cells”. In: *Current Opinion in Biotechnology*. Chemical Biotechnology • Pharmaceutical Biotechnology 48 (2017), pp. 61–68. DOI: 10.1016/j.copbio.2017.03.012.

[67] B. Albert Griffin, Stephen R. Adams, and Roger Y. Tsien. “Specific Covalent Labeling of Recombinant Protein Molecules Inside Live Cells”. In: *Science* (1998). DOI: 10.1126/science.281.5374.269.

[68] Lawrence W. Miller et al. “In Vivo Protein Labeling with Trimethoprim Conjugates: A Flexible Chemical Tag”. In: *Nature Methods* 2.4 (2005), pp. 255–257. DOI: 10.1038/nmeth749.

[69] Sarah S. Gallagher et al. “An In Vivo Covalent TMP-Tag Based on Proximity-Induced Reactivity”. In: *ACS Chemical Biology* 4.7 (2009), pp. 547–556. DOI: 10.1021/cb900062k.

[70] Jonas Wilhelm et al. “Kinetic and Structural Characterization of the Self-Labeling Protein Tags HaloTag7, SNAP-tag, and CLIP-tag”. In: *Biochemistry* (2021). DOI: 10.1021/acs.biochem.1c00258.

[71] Antje Keppler et al. “A General Method for the Covalent Labeling of Fusion Proteins with Small Molecules in Vivo”. In: *Nature Biotechnology* 21.1 (2003), pp. 86–89. DOI: 10.1038/nbt765.

[72] Thomas Gronemeyer et al. “Directed Evolution of O6-alkylguanine-DNA Alkyltransferase for Applications in Protein Labeling”. In: *Protein Engineering, Design and Selection* 19.7 (2006), pp. 309–316. DOI: 10.1093/protein/gzl014.

[73] Arnaud Gautier et al. “An Engineered Protein Tag for Multiprotein Labeling in Living Cells”. In: *Chemistry & Biology* 15.2 (2008), pp. 128–136. DOI: 10.1016/j.chembiol.2008.01.007.

[74] Georgyi V. Los et al. “HaloTag: A Novel Protein Labeling Technology for Cell Imaging and Protein Analysis”. In: *ACS Chemical Biology* 3.6 (2008), pp. 373–382. DOI: 10.1021/cb800025k.

[75] Jiří Damborský and Jaroslav Koča. “Analysis of the Reaction Mechanism and Substrate Specificity of Haloalkane Dehalogenases by Sequential and Structural Comparisons”. In: *Protein Engineering, Design and Selection* 12.11 (1999), pp. 989–998. DOI: 10.1093/protein/12.11.989.

[76] Lance P. Encell et al. “Development of a Dehalogenase-Based Protein Fusion Tag Capable of Rapid, Selective and Covalent Attachment to Customizable Ligands”. In: *Current Chemical Genomics* 6.1 (2012). DOI: 10.2174/1875397301206010055.

[77] Gražvydas Lukinavičius et al. “A Near-Infrared Fluorophore for Live-Cell Super-Resolution Microscopy of Cellular Proteins”. In: *Nature Chemistry* 5.2 (2013), pp. 132–139. DOI: 10.1038/nchem.1546.

[78] Kirsten Deprey and Joshua A. Kritzer. “HaloTag Forms an Intramolecular Disulfide”. In: *Bioconjugate Chemistry* 32.5 (2021), pp. 964–970. DOI: 10.1021/acs.bioconjchem.1c00113.

[79] Michelle S. Frei et al. “Engineered HaloTag Variants for Fluorescence Lifetime Multiplexing”. In: *Nature Methods* (2021), pp. 1–6. DOI: 10.1038/s41592-021-01341-x.

[80] Georgeta Crivat and Justin W. Taraska. “Imaging Proteins inside Cells with Fluorescent Tags”. In: *Trends in Biotechnology* 30.1 (2012), pp. 8–16. DOI: 10.1016/j.tibtech.2011.08.002.

[81] David Virant et al. “A Peptide Tag-Specific Nanobody Enables High-Quality Labeling for dSTORM Imaging”. In: *Nature Communications* 9.1 (2018), p. 930. DOI: 10.1038/s41467-018-03191-2.

[82] Hansjörg Götzke et al. “The ALFA-tag Is a Highly Versatile Tool for Nanobody-Based Bioscience Applications”. In: *Nature Communications* 10.1 (1 2019), p. 4403. DOI: 10.1038/s41467-019-12301-7.

[83] Kathrin Lang and Jason W. Chin. “Cellular Incorporation of Unnatural Amino Acids and Bioorthogonal Labeling of Proteins”. In: *Chemical Reviews* 114.9 (2014), pp. 4764–4806. DOI: 10.1021/cr400355w.

[84] Samuel L. Scinto et al. “Bioorthogonal Chemistry”. In: *Nature Reviews Methods Primers* 1.1 (1 2021), pp. 1–23. DOI: 10.1038/s43586-021-00028-z.

[85] Melissa L. Blackman, Maksim Royzen, and Joseph M. Fox. “Tetrazine Ligation: Fast Bioconjugation Based on Inverse-Electron-Demand Diels-Alder Reactivity”. In: *Journal of the American Chemical Society* 130.41 (2008), pp. 13518–13519. DOI: 10.1021/ja8053805.

[86] B.L. Oliveira, Z. Guo, and G. J.L. Bernardes. “Inverse Electron Demand Diels–Alder Reactions in Chemical Biology”. In: *Chemical Society Reviews* 46.16 (2017), pp. 4895–4950. DOI: 10.1039/C7CS00184C.

[87] Susanne Mayer and Kathrin Lang. “Tetrazines in Inverse-Electron-Demand Diels–Alder Cycloadditions and Their Use in Biology”. In: *Synthesis* 49.4 (2017), pp. 830–848. DOI: 10.1055/s-0036-1588682.

[88] M W Wessendorf, S J Tallaksen-Greene, and R M Wohlhueter. “A Spectrophotometric Method for Determination of Fluorophore-to-Protein Ratios in Conjugates of the Blue Fluorophore 7-Amino-4-Methylcoumarin-3-Acetic Acid (AMCA”. In: *The journal of histochemistry and cytochemistry* 38.1 (1990), pp. 87–94. DOI: 10.1177/38.1.1688452.

[89] Xiaoli Sun et al. “Development of SNAP-Tag Fluorogenic Probes for Wash-Free Fluorescence Imaging”. In: *ChemBioChem* 12.14 (2011), pp. 2217–2226. DOI: 10.1002/cbic.201100173.

[90] Davide Calebiro et al. “Single-Molecule Analysis of Fluorescently Labeled G-protein–Coupled Receptors Reveals Complexes with Distinct Dynamics and Organization”. In: *Proceedings of the National Academy of Sciences* 110.2 (2013), pp. 743–748. DOI: 10.1073/pnas.1205798110.

[91] Stephan Wilmes et al. “Receptor Dimerization Dynamics as a Regulatory Valve for Plasticity of Type I Interferon Signaling”. In: *Journal of Cell Biology* 209.4 (2015), pp. 579–593. DOI: 10.1083/jcb.201412049.

[92] Jervis Vermal Thevathasan et al. “Nuclear Pores as Versatile Reference Standards for Quantitative Superresolution Microscopy”. In: *Nature Methods* 16.10 (2019), pp. 1045–1053. DOI: 10.1038/s41592-019-0574-9.

[93] Yi-Chi C. Liu et al. “Specific and Reversible Immobilization of Histidine-Tagged Proteins on Functionalized Silicon Nanowires”. In: *Nanotechnology* 21.24 (2010), p. 245105. DOI: 10.1088/0957-4484/21/24/245105.

[94] Friedrich Roder et al. “Rapid Transfer of Transmembrane Proteins for Single Molecule Dimerization Assays in Polymer-Supported Membranes”. In: *ACS Chemical Biology* 9.11 (2014), pp. 2479–2484. DOI: 10.1021/cb5005806.

[95] Peter J. Bosch et al. “Evaluation of Fluorophores to Label SNAP-Tag Fused Proteins for Multicolor Single-Molecule Tracking Microscopy in Live Cells”. In: *Biophysical Journal* 107.4 (2014), pp. 803–814. DOI: 10.1016/j.bpj.2014.06.040.

[96] Valerie C. Coffman and Jian-Qiu Wu. “Counting Protein Molecules Using Quantitative Fluorescence Microscopy”. In: *Trends in Biochemical Sciences* 37.11 (2012), pp. 499–506. DOI: 10.1016/j.tibs.2012.08.002.

[97] Johan Hummert, Stanimir Asenov Tashev, and Dirk-Peter Herten. “An Update on Molecular Counting in Fluorescence Microscopy”. In: *The International Journal of Biochemistry & Cell Biology* 135 (2021), p. 105978. DOI: 10.1016/j.biocel.2021.105978.

[98] Michelle A. Digman et al. “Mapping the Number of Molecules and Brightness in the Laser Scanning Microscope”. In: *Biophysical Journal* 94.6 (2008), pp. 2320–2332. DOI: 10.1529/biophysj.107.114645.

[99] Molly J. Rossow et al. “Raster Image Correlation Spectroscopy in Live Cells”. In: *Nature Protocols* 5.11 (11 2010), pp. 1761–1774. DOI: 10.1038/nprot.2010.122.

[100] Michael R. Stoneman et al. “A General Method to Quantify Ligand-Driven Oligomerization from Fluorescence-Based Images”. In: *Nature Methods* 16.6 (6 2019), pp. 493–496. DOI: 10.1038/s41592-019-0408-9.

[101] Jonas Ries and Petra Schwille. “Fluorescence Correlation Spectroscopy”. In: *BioEssays* 34.5 (2012), pp. 361–368. DOI: 10.1002/bies.201100111.

[102] Paul W. Wiseman. “Image Correlation Spectroscopy: Principles and Applications”. In: *Cold Spring Harbor Protocols* 2015.4 (2015), pdb.top086124. DOI: 10.1101/pdb.top086124.

[103] Lan Yu et al. “A Comprehensive Review of Fluorescence Correlation Spectroscopy”. In: *Frontiers in Physics* 9 (2021), p. 110. DOI: 10.3389/fphy.2021.644450.

[104] D. Gross and W. W. Webb. “Molecular Counting of Low-Density Lipoprotein Particles as Individuals and Small Clusters on Cell Surfaces”. In: *Biophysical Journal* 49.4 (1986), pp. 901–911. DOI: 10.1016/S0006-3495(86)83718-3.

[105] Thomas Schmidt et al. “Local Stoichiometries Determined by Counting Individual Molecules”. In: *Analytical Chemistry* 68.24 (1996), pp. 4397–4401. DOI: 10.1021/ac960710g.

[106] Sarah A. Mutch et al. “Deconvolving Single-Molecule Intensity Distributions for Quantitative Microscopy Measurements”. In: *Biophysical Journal* 92.8 (2007), pp. 2926–2943. DOI: 10.1529/biophysj.106.101428.

[107] Laurent Cognet et al. “Robust Single-Molecule Approach for Counting Autofluorescent Proteins”. In: *Journal of Biomedical Optics* 13.3 (2008), p. 031216. DOI: 10.1117/1.2940600.

[108] Ute Resch-Genger et al. “Quantum Dots versus Organic Dyes as Fluorescent Labels”. In: *Nature Methods* 5.9 (2008), pp. 763–775. DOI: 10.1038/nmeth.1248.

[109] Jürgen J. Schmied et al. “Fluorescence and Super-Resolution Standards Based on DNA Origami”. In: *Nature Methods* 9.12 (2012), pp. 1133–1134. DOI: 10.1038/nmeth.2254.

[110] Hermann J. Gruber et al. “Anomalous Fluorescence Enhancement of Cy3 and Cy3.5 versus Anomalous Fluorescence Loss of Cy5 and Cy7 upon Covalent Linking to IgG and Noncovalent Binding to Avidin”. In: *Bioconjugate Chemistry* 11.5 (2000), pp. 696–704. DOI: 10.1021/bc000015m.

[111] Malea Kneen et al. “Green Fluorescent Protein as a Noninvasive Intracellular pH Indicator”. In: *Biophysical Journal* 74.3 (1998), pp. 1591–1599. DOI: 10.1016/S0006-3495(98)77870-1.

[112] Ronald D. Vale et al. “Direct Observation of Single Kinesin Molecules Moving along Microtubules”. In: *Nature* 380.6573 (6573 1996), pp. 451–453. DOI: 10.1038/380451a0.

[113] Maximilian H. Ulbrich and Ehud Y. Isacoff. “Subunit Counting in Membrane-Bound Proteins”. In: *Nature Methods* 4.4 (2007), pp. 319–321. DOI: 10.1038/nmeth1024.

[114] Somes K. Das et al. “Membrane Protein Stoichiometry Determined from the Step-Wise Photobleaching of Dye-Labelled Subunits”. In: *Chembiochem: A European Journal of Chemical Biology* 8.9 (2007), pp. 994–999. DOI: 10.1002/cbic.200600474.

[115] Yalei Chen et al. “Molecular Counting by Photobleaching in Protein Complexes with Many Subunits: Best Practices and Application to the Cellulose Synthesis Complex”. In: *Molecular Biology of the Cell* 25.22 (2014), pp. 3630–3642. DOI: 10.1091/mbc.e14-06-1146.

[116] Rose E. Dixon et al. “Graded Ca2+/Calmodulin-Dependent Coupling of Voltage-Gated CaV1. 2 Channels”. In: *Elife* 4 (2015), e05608.

[117] Ryan J. Arant and Maximilian H. Ulbrich. “Deciphering the Subunit Composition of Multimeric Proteins by Counting Photobleaching Steps”. In: *ChemPhysChem* 15.4 (2014), pp. 600–605. DOI: 10.1002/cphc.201301092.

[118] Benjamin D. Engel, William B. Ludington, and Wallace F. Marshall. “Intraflagellar Transport Particle Size Scales Inversely with Flagellar Length: Revisiting the Balance-Point Length Control Model”. In: *Journal of Cell Biology* 187.1 (2009), pp. 81–89. DOI: 10.1083/jcb.200812084.

[119] Jinghe Yuan et al. “Analysis of the Steps in Single-Molecule Photobleaching Traces by Using the Hidden Markov Model and Maximum-Likelihood Clustering”. In: *Chemistry - An Asian Journal* 9.8 (2014), pp. 2303–2308. DOI: 10.1002/asia.201402147.

[120] Lucas P. Watkins and Haw Yang. “Detection of Intensity Change Points in Time-Resolved Single-Molecule Measurements”. In: *The Journal of Physical Chemistry B* 109.1 (2005), pp. 617–628. DOI: 10.1021/jp0467548.

[121] Rabindra V. Shivnaraine et al. “Single-Molecule Analysis of the Supramolecular Organization of the M2 Muscarinic Receptor and the Gαi1 Protein”. In: *Journal of the American Chemical Society* 138.36 (2016), pp. 11583–11598. DOI: 10.1021/jacs.6b04032.

[122] Hugo McGuire et al. “Automating Single Subunit Counting of Membrane Proteins in Mammalian Cells”. In: *Journal of Biological Chemistry* 287.43 (2012), pp. 35912–35921. DOI: 10.1074/jbc.M112.402057.

[123] Konstantinos Tsekouras et al. “A Novel Method to Accurately Locate and Count Large Numbers of Steps by Photobleaching”. In: *Molecular Biology of the Cell* 27.22 (2016), pp. 3601–3615. DOI: 10.1091/mbc.e16-06-0404.

[124] Jon Garry et al. “Bayesian Counting of Photobleaching Steps with Physical Priors”. In: *The Journal of Chemical Physics* 152.2 (2020), p. 024110. DOI: 10.1063/1.5132957.

[125] Th. Basché and W. E. Moerner. “Photon Antibunching in the Fluorescence of a Single Dye Molecule Trapped in a Solid”. In: *Physical Review Letters* 69.10 (1992), pp. 1516–1519. DOI: 10.1103/PhysRevLett.69.1516.

[126] Kenneth D. Weston et al. “Measuring the Number of Independent Emitters in Single-Molecule Fluorescence Images and Trajectories Using Coincident Photons”. In: *Analytical Chemistry* 74.20 (2002), pp. 5342–5349. DOI: 10.1021/ac025730z.

[127] Kristin S. Grußmayer and Dirk-Peter Herten. “Photon Antibunching in Single Molecule Fluorescence Spectroscopy”. In: *Advanced Photon Counting*. Springer Series on Fluorescence. Springer, 2014, pp. 159–190. ISBN: 978-3-319-15635-4 978-3-319-15636-1. DOI: 10.1007/4243_2014_71.

[128] Kristin S. Grußmayer et al. “Differentiation between Shallow and Deep Charge Trap States on Single Poly(3-Hexylthiophene) Chains through Fluorescence Photon Statistics”. In: *ChemPhysChem* 16.17 (2015), pp. 3578–3583. DOI: 10.1002/cphc.201500719.

[129] K. S. Grußmayer and D.-P. Herten. “Time-Resolved Molecule Counting by Photon Statistics across the Visible Spectrum”. In: *Physical Chemistry Chemical Physics* 19.13 (2017), pp. 8962–8969. DOI: 10.1039/C7CP00363C.

[130] Anton Kurz et al. “Counting Fluorescent Dye Molecules on DNA Origami by Means of Photon Statistics”. In: *Small* 9.23 (2013), pp. 4061–4068. DOI: 10.1002/smll.201300619.

[131] Haisen Ta, Jürgen Wolfrum, and Dirk-Peter Herten. “An Extended Scheme for Counting Fluorescent Molecules by Photon-Antibunching”. In: *Laser Physics* 20.1 (2010), pp. 119–124. DOI: 10.1134/S1054660X09170204.

[132] Haisen Ta et al. “Mapping Molecules in Scanning Far-Field Fluorescence Nanoscopy”. In: *Nature Communications* 6 (2015), p. 7977. DOI: 10.1038/ncomms8977.

[133] H. Deschout et al. “Progress in Quantitative Single-Molecule Localization Microscopy”. In: *Histochemistry and Cell Biology* 142.1 (2014), pp. 5–17. DOI: 10.1007/s00418-014-1217-y.

[134] Paolo Annibale et al. “Quantitative Photo Activated Localization Microscopy: Unraveling the Effects of Photoblinking”. In: *PloS One* 6.7 (2011), e22678. DOI: 10.1371/journal.pone.0022678.

[135] Sang-Hyuk Lee et al. “Counting Single Photoactivatable Fluorescent Molecules by Photoactivated Localization Microscopy (PALM)”. In: *Proceedings of the National Academy of Sciences* 109.43 (2012), pp. 17436–17441. DOI: 10.1073/pnas.1215175109.

[136] Elias M. Puchner et al. “Counting Molecules in Single Organelles with Superresolution Microscopy Allows Tracking of the Endosome Maturation Trajectory”. In: *Proceedings of the National Academy of Sciences of the United States of America* 110.40 (2013), pp. 16015–16020. DOI: 10.1073/pnas.1309676110.

[137] Santosh Adhikari, Joe Moscatelli, and Elias M. Puchner. “Quantitative Live-Cell PALM Reveals Nanoscopic Faa4 Redistributions and Dynamics on Lipid Droplets during Metabolic Transitions of Yeast”. In: *Molecular Biology of the Cell* 32.17 (2021), pp. 1565–1578. DOI: 10.1091/mbc.E20-11-0695.

[138] Nadine Ehmann et al. “Quantitative Super-Resolution Imaging of Bruchpilot Distinguishes Active Zone States”. In: *Nature Communications* 5 (2014), p. 4650. DOI: 10.1038/ncomms5650.

[139] Christos Karathanasis et al. “Molecule Counts in Localization Microscopy with Organic Fluorophores”. In: *ChemPhysChem* 18.8 (2017), pp. 942–948. DOI: 10.1002/cphc.201601425.

[140] Gerhard Hummer, Franziska Fricke, and Mike Heilemann. “Model-Independent Counting of Molecules in Single-Molecule Localization Microscopy”. In: *Molecular Biology of the Cell* 27.22 (2016), pp. 3637–3644. DOI: 10.1091/mbc.e16-07-0525.

[141] Ralf Jungmann et al. “Multiplexed 3D Cellular Super-Resolution Imaging with DNA-PAINT and Exchange-PAINT”. In: *Nature Methods* 11.3 (2014), pp. 313–318. DOI: 10.1038/nmeth.2835.

[142] Ralf Jungmann et al. “Quantitative Super-Resolution Imaging with qPAINT”. In: *Nature Methods* 13.5 (5 2016), pp. 439–442. DOI: 10.1038/nmeth.3804.

[143] Johannes Stein et al. “Towards Absolute Molecular Numbers in DNA-PAINT”. In: *Nano Letters* (2019). DOI: 10.1021/acs.nanolett.9b03546.

[144] Isuru Jayasinghe et al. “True Molecular Scale Visualization of Variable Clustering Properties of Ryanodine Receptors”. In: *Cell Reports* 22.2 (2018), pp. 557–567. DOI: 10.1016/j.celrep.2017.12.045.

[145] Daniel J. Nieves et al. “tagPAINT: Stoichiometric and Covalent Labelling of Protein Tags for Multiplexed and Quantitative DNA-PAINT Imaging”. In: *bioRxiv* (2019), p. 604462. DOI: 10.1101/604462.

[146] Megan D. Joseph et al. “Quantitative Super-Resolution Imaging for the Analysis of GPCR Oligomerization”. In: *Biomolecules* 11.10 (10 2021), p. 1503. DOI: 10.3390/biom11101503.

[147] Johannes Stein et al. “Calibration Free Counting of Low Molecular Copy Numbers in Single DNA-PAINT Localization Clusters”. In: *bioRxiv* (2021), p. 2021.08.17.456678. DOI: 10.1101/2021.08.17.456678.

[148] Jennifer C. Waters. “Accuracy and Precision in Quantitative Fluorescence Microscopy”. In: *Journal of Cell Biology* 185.7 (2009), pp. 1135–1148. DOI: 10.1083/jcb.200903097.

[149] Paula Montero Llopis et al. “Best Practices and Tools for Reporting Reproducible Fluorescence Microscopy Methods”. In: *Nature Methods* 18.12 (12 2021), pp. 1463–1476. DOI: 10.1038/s41592-021-01156-w.

[150] Melissa Linkert et al. “Metadata Matters: Access to Image Data in the Real World”. In: *Journal of Cell Biology* 189.5 (2010), pp. 777–782. DOI: 10.1083/jcb.201004104.

[151] Mathias Hammer et al. “Towards Community-Driven Metadata Standards for Light Microscopy: Tiered Specifications Extending the OME Model”. In: *Nature Methods* 18.12 (12 2021), pp. 1427–1440. DOI: 10.1038/s41592-021-01327-9.

[152] Susanne Kunis et al. “MDEmic: A Metadata Annotation Tool to Facilitate Management of FAIR Image Data in the Bioimaging Community”. In: *Nature Methods* 18.12 (12 2021), pp. 1416–1417. DOI: 10.1038/s41592-021-01288-z.

[153] Patrick Theer, Cyril Mongis, and Michael Knop. “PSFj: Know Your Fluorescence Microscope”. In: *Nature Methods* 11.10 (10 2014), pp. 981–982. DOI: 10.1038/nmeth.3102.

[154] Sungyong You et al. “Microscope Calibration Protocol for Single-Molecule Microscopy”. In: *Optics Express* 29.1 (2021), pp. 182–207. DOI: 10.1364/OE.408361.

[155] Markus Mund and Jonas Ries. “How Good Are My Data? Reference Standards in Superresolution Microscopy”. In: *Molecular Biology of the Cell* 31.19 (2020), pp. 2093–2096. DOI: 10.1091/mbc.E19-04-0189.

[156] Christian Steinhauer et al. “DNA Origami as a Nanoscopic Ruler for Super-Resolution Microscopy”. In: *Angewandte Chemie International Edition* 48.47 (2009), pp. 8870–8873. DOI: 10.1002/anie.200903308.

[157] Eric F. Pettersen et al. “UCSF Chimera—A Visualization System for Exploratory Research and Analysis - Pettersen”. In: *Journal of Computational Chemistry* 25.25 (2004), pp. 1605–1612. DOI: DOI10.1002/jcc.20084.

[158] Do-Nyun Kim et al. “Quantitative Prediction of 3D Solution Shape and Flexibility of Nucleic Acid Nanostructures”. In: *Nucleic Acids Research* 40.7 (2012), pp. 2862–2868. DOI: 10.1093/nar/gkr1173.

[159] Clarissa Liesche et al. “Automated Analysis of Single-Molecule Photobleaching Data by Statistical Modeling of Spot Populations”. In: *Biophysical Journal* 109.11 (2015), pp. 2352–2362. DOI: 10.1016/j.bpj.2015.10.035.

[160] John M. Murray. “An icosahedral virus as a fluorescent calibration standard: a method for counting protein molecules in cells by fluorescence microscopy”. In: *Journal of Microscopy* 267.2 (2017), pp. 193–213. DOI: 10.1111/jmi.12559.

[161] Francesca Cella Zanacchi et al. “A DNA Origami Platform for Quantifying Protein Copy Number in Super-Resolution”. In: *Nature Methods* 14.8 (2017), pp. 789–792. DOI: 10.1038/nmeth.4342.

[162] Paul W. K. Rothemund. “Folding DNA to Create Nanoscale Shapes and Patterns”. In: *Nature* 440.7082 (2006), pp. 297–302. DOI: 10.1038/nature04586.

[163] Jürgen J. Schmied et al. “DNA Origami–Based Standards for Quantitative Fluorescence Microscopy”. In: *Nature Protocols* 9.6 (2014), pp. 1367–1391. DOI: 10.1038/nprot.2014.079.

[164] Maximilian T. Strauss et al. “Quantifying Absolute Addressability in DNA Origami with Molecular Resolution”. In: *Nature Communications* 9.1 (2018), p. 1600. DOI: 10.1038/s41467-018-04031-z.

[165] Kevin Jahnke et al. “Choice of Fluorophore Affects Dynamic DNA Nanostructures”. In: *Nucleic Acids Research* 49.7 (2021), pp. 4186–4195. DOI: 10.1093/nar/gkab201.

[166] Nela Durisic et al. “Stoichiometry of the Human Glycine Receptor Revealed by Direct Subunit Counting”. In: *Journal of Neuroscience* 32.37 (2012), pp. 12915–12920. DOI: 10.1523/JNEUROSCI.2050-12.2012.

[167] Nela Durisic et al. “Single-Molecule Evaluation of Fluorescent Protein Photoactivation Efficiency Using an in Vivo Nanotemplate”. In: *Nature Methods* 11.2 (2 2014), pp. 156–162. DOI: 10.1038/nmeth.2784.

[168] Kieran Finan, Anika Raulf, and Mike Heilemann. “A Set of Homo-Oligomeric Standards Allows Accurate Protein Counting”. In: *Angewandte Chemie International Edition* 54.41 (2015), pp. 12049–12052. DOI: 10.1002/anie.201505664.

[169] M. Delarue et al. “mTORC1 Controls Phase Separation and the Biophysical Properties of the Cytoplasm by Tuning Crowding”. In: *Cell* 174.2 (2018), 338–349.e20. DOI: 10.1016/j.cell.2018.05.042.

[170] Fusamichi Akita et al. “The Crystal Structure of a Virus-like Particle from the Hyperthermophilic Archaeon Pyrococcus Furiosus Provides Insight into the Evolution of Viruses”. In: *Journal of Molecular Biology* 368.5 (2007), pp. 1469–1483. DOI: 10.1016/j.jmb.2007.02.075.

[171] Ryoichi Arai et al. “Design of the Linkers Which Effectively Separate Domains of a Bifunctional Fusion Protein”. In: *Protein Engineering, Design and Selection* 14.8 (2001), pp. 529–532. DOI: 10.1093/protein/14.8.529.

[172] Marilyn D. Resh. “Fatty Acylation of Proteins: New Insights into Membrane Targeting of Myristoylated and Palmitoylated Proteins”. In: *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1451.1 (1999), pp. 1–16. DOI: 10.1016/S0167-4889(99)00075-0.

[173] Susan Swift et al. “Rapid Production of Retroviruses for Efficient Gene Delivery to Mammalian Cells Using 293T Cell–Based Systems”. In: *Current Protocols in Immunology* 31.1 (1999), pp. 10.17.14–10.17.29. DOI: 10.1002/0471142735.im1017cs31.

[174] Siegfried Hänselmann. *Quantitative Microscopy: Measuring Membrane Receptor Interactions in Live Cells*. Dissertation, Ruperto-Carola University of Heidelberg, 2017.

[175] Daniel Sage et al. “Super-Resolution Fight Club: Assessment of 2D and 3D Single-Molecule Localization Microscopy Software”. In: *Nature Methods* (2019), p. 1. DOI: 10.1038/s41592-019-0364-4.

[176] Khuloud Jaqaman et al. “Robust Single-Particle Tracking in Live-Cell Time-Lapse Sequences”. In: *Nature Methods* 5.8 (8 2008), pp. 695–702. DOI: 10.1038/nmeth.1237.

[177] Arthur D Edelstein et al. “Advanced Methods of Microscope Control Using μManager Software”. In: *Journal of Biological Methods* 1.2 (2014), p. 10. DOI: 10.14440/jbm.2014.36.

[178] Martin Ovesný et al. “ThunderSTORM: A Comprehensive ImageJ Plug-in for PALM and STORM Data Analysis and Super-Resolution Imaging”. In: *Bioinformatics* 30.16 (2014), pp. 2389–2390. DOI: 10.1093/bioinformatics/btu202.

[179] Tibor Novák et al. “TestSTORM: Versatile Simulator Software for Multimodal Super-Resolution Localization Fluorescence Microscopy”. In: *Scientific Reports* 7.1 (1 2017), p. 951. DOI: 10.1038/s41598-017-01122-7.

[180] Ruby M. Phelps et al. “NCI-navy Medical Oncology Branch Cell Line Data Base”. In: *Journal of Cellular Biochemistry* 63.S24 (1996), pp. 32–91. DOI: 10.1002/jcb.240630505.

[181] Keril J. Blight, Jane A. McKeating, and Charles M. Rice. “Highly Permissive Cell Lines for Subgenomic and Genomic Hepatitis C Virus RNA Replication”. In: *Journal of Virology* 76.24 (2002), pp. 13001–13014. DOI: 10.1128/JVI.76.24.13001-13014.2002.

[182] L. Zhao et al. “Intracellular Water-Specific MR of Microbead-Adherent Cells: The HeLa Cell Intracellular Water Exchange Lifetime”. In: *NMR in Biomedicine* 21.2 (2008), pp. 159–164. DOI: 10.1002/nbm.1173.

[183] Jonathan B. Grimm et al. “A General Method to Improve Fluorophores for Live-Cell and Single-Molecule Microscopy”. In: *Nature Methods* 12.3 (2015), pp. 244–250. DOI: 10.1038/nmeth.3256.

[184] Graham T. Dempsey et al. “Evaluation of Fluorophores for Optimal Performance in Localization-Based Super-Resolution Imaging”. In: *Nature Methods* 8.12 (2011), pp. 1027–1036. DOI: 10.1038/nmeth.1768.

[185] Yasser Gidi et al. “Unifying Mechanism for Thiol-Induced Photoswitching and Photostability of Cyanine Dyes”. In: *Journal of the American Chemical Society* 142.29 (2020), pp. 12681–12689. DOI: 10.1021/jacs.0c03786.

[186] Junhong Min et al. “FALCON: Fast and Unbiased Reconstruction of High-Density Super-Resolution Microscopy Data”. In: *Scientific Reports* 4.1 (1 2014), p. 4577. DOI: 10.1038/srep04577.

[187] Elias Nehme et al. “Deep-STORM: Super-Resolution Single-Molecule Microscopy by Deep Learning”. In: *Optica* 5.4 (2018), pp. 458–464. DOI: 10.1364/OPTICA.5.000458.

[188] Artur Speiser et al. “Deep Learning Enables Fast and Dense Single-Molecule Localization with High Accuracy”. In: *Nature Methods* 18.9 (9 2021), pp. 1082–1090. DOI: 10.1038/s41592-021-01236-x.

[189] Sarah L. Latty et al. “Referenced Single-Molecule Measurements Differentiate between GPCR Oligomerization States”. In: *Biophysical Journal* 109.9 (2015), pp. 1798–1806. DOI: 10.1016/j.bpj.2015.09.004.

[190] Martin Oheim et al. “Calibrating Evanescent-Wave Penetration Depths for Biological TIRF Microscopy”. In: *Biophysical Journal* 117.5 (2019), pp. 795–809. DOI: 10.1016/j.bpj.2019.07.048.

[191] Thomas P. Burghardt. “Measuring Incidence Angle for Through-the-Objective Total Internal Reflection Fluorescence Microscopy”. In: *Journal of Biomedical Optics* 17.12 (2012), p. 126007. DOI: 10.1117/1.JBO.17.12.126007.

[192] R. Fiolka et al. “Even Illumination in Total Internal Reflection Fluorescence Microscopy Using Laser Light”. In: *Microscopy Research and Technique* 71.1 (2008), pp. 45–50. DOI: 10.1002/jemt.20527.

[193] Alexa L. Mattheyses, Keith Shaw, and Daniel Axelrod. “Effective Elimination of Laser Interference Fringing in Fluorescence Microscopy by Spinning Azimuthal Incidence Angle”. In: *Microscopy Research and Technique* 69.8 (2006), pp. 642–647. DOI: 10.1002/jemt.20334.

[194] Masato Yasui et al. “Automated Single-Molecule Imaging in Living Cells”. In: *Nature Communications* 9.1 (1 2018), p. 3061. DOI: 10.1038/s41467-018-05524-7.

[195] Johan Hummert et al. “Photobleaching Step Analysis for Robust Determination of Protein Complex Stoichiometries”. In: *Molecular Biology of the Cell* 32.21 (2021), ar35. DOI: 10.1091/mbc.E20-09-0568.

[196] Brian C. Carter, Michael Vershinin, and Steven P. Gross. “A Comparison of Step-Detection Methods: How Well Can You Do?” In: *Biophysical Journal* 94.1 (2008), pp. 306–319. DOI: 10.1529/biophysj.107.110601.

[197] Bennett Kalafut and Koen Visscher. “An Objective, Model-Independent Method for Detection of Non-Uniform Steps in Noisy Signals”. In: *Computer Physics Communications* 179.10 (2008), pp. 716–723. DOI: 10.1016/j.cpc.2008.06.008.

[198] Kristin Stefanie Grußmayer. *Fluorescence Quantification by Photon Statistics: From Objective Characterization to Application*. Dissertation, Ruperto-Carola University of Heidelberg, 2015. DOI: 10.11588/heidok.00020073.

[199] S. Walter Englander, Dorothy B. Calhoun, and Joan J. Englander. “Biochemistry without Oxygen”. In: *Analytical Biochemistry* 161.2 (1987), pp. 300–306. DOI: 10.1016/0003-2697(87)90454-4.

[200] P. V. Patil and D. P. Ballou. “The Use of Protocatechuate Dioxygenase for Maintaining Anaerobic Conditions in Biochemical Experiments”. In: *Analytical Biochemistry* 286.2 (2000), pp. 187–192. DOI: 10.1006/abio.2000.4802.

[201] Tobias MP Hartwich et al. “A Stable, High Refractive Index, Switching Buffer for Super-Resolution Imaging”. In: *bioRxiv* (2018), p. 465492. DOI: 10.1101/465492.

[202] S. Yagi and H. Inoue. “The Absorption of Oxygen into Sodium Sulphite Solution”. In: *Chemical Engineering Science* 17.6 (1962), pp. 411–421. DOI: 10.1016/0009-2509(62)85010-6.

[203] L Song et al. “Photobleaching Kinetics of Fluorescein in Quantitative Fluorescence Microscopy.” In: *Biophysical Journal* 68.6 (1995), pp. 2588–2600. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1282169/ (visited on 01/07/2022).

[204] C. Eggeling et al. “Photobleaching of Fluorescent Dyes under Conditions Used for Single-Molecule Detection: Evidence of Two-Step Photolysis”. In: *Analytical Chemistry* 70.13 (1998), pp. 2651–2659. DOI: 10.1021/ac980027p.

[205] Christian A. Wurm et al. “Novel Red Fluorophores with Superior Performance in STED Microscopy”. In: *Optical Nanoscopy* 1.1 (2012), p. 7. DOI: 10.1186/2192-2853-1-7.

[206] Nathan C. Shaner et al. “A Bright Monomeric Green Fluorescent Protein Derived from Branchiostoma Lanceolatum”. In: *Nature Methods* 10.5 (2013), pp. 407–409. DOI: 10.1038/nmeth.2413.

[207] Roman S. Erdmann et al. “Labeling Strategies Matter for Super-Resolution Microscopy: A Comparison between HaloTags and SNAP-Tags”. In: *Cell Chemical Biology* 26.4 (2019), 584–592.e6. DOI: 10.1016/j.chembiol.2019.01.003.

[208] Simon Sehayek et al. “A High-Throughput Image Correlation Method for Rapid Analysis of Fluorophore Photoblinking and Photobleaching Rates”. In: *ACS Nano* 13.10 (2019), pp. 11955–11966. DOI: 10.1021/acsnano.9b06033.

[209] Harindarpal S. Gill and David Eisenberg. “The Crystal Structure of Phosphinothricin in the Active Site of Glutamine Synthetase Illuminates the Mechanism of Enzymatic Inhibition”. In: *Biochemistry* 40.7 (2001), pp. 1903–1912. DOI: 10.1021/bi002438h.

[210] Yiming Li et al. “Real-Time 3D Single-Molecule Localization Using Experimental Point Spread Functions”. In: *Nature Methods* 15.5 (5 2018), pp. 367–369. DOI: 10.1038/nmeth.4661.

[211] Douglas G. Bonett. “Confidence Interval for a Coefficient of Quartile Variation”. In: *Computational Statistics & Data Analysis* 50.11 (2006), pp. 2953–2957. DOI: 10.1016/j.csda.2005.05.007.

[212] Alexander von Appen et al. “In Situ Structural Analysis of the Human Nuclear Pore Complex”. In: *Nature* 526.7571 (2015), pp. 140–143. DOI: 10.1038/nature15381.

[213] Qinsi Zheng et al. “Ultra-Stable Organic Fluorophores for Single-Molecule Research”. In: *Chemical Society Reviews* 43.4 (2014), pp. 1044–1056. DOI: 10.1039/C3CS60237K.

[214] Roger B. Altman et al. “Cyanine Fluorophore Derivatives with Enhanced Photostability”. In: *Nature Methods* 9.1 (2012), pp. 68–71. DOI: 10.1038/nmeth.1774.

[215] Jochem H. Smit et al. “On the Impact of Competing Intra- and Intermolecular Triplet-State Quenching on Photobleaching and Photoswitching Kinetics of Organic Fluorophores”. In: *Physical Chemistry Chemical Physics* 21.7 (2019), pp. 3721–3733. DOI: 10.1039/C8CP05063E.

[216] Thomas Gensch, Martin Böhmer, and Pedro F. Aramendía. “Single Molecule Blinking and Photobleaching Separated by Wide-Field Fluorescence Microscopy”. In: *The Journal of Physical Chemistry A* 109.30 (2005), pp. 6652–6658. DOI: 10.1021/jp0510847.

[217] Sangmin Ji et al. “Investigation on Photobleaching of Fluorophores: Effect of Excitation Power and Buffer System”. In: *Bulletin of the Korean Chemical Society* n/a.n/a (2021). DOI: 10.1002/bkcs.12455.

[218] U Kubitscheck et al. “Single Nuclear Pores Visualized by Confocal Microscopy and Image Processing.” In: *Biophysical Journal* 70.5 (1996), pp. 2067–2077. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1225182/ (visited on 01/10/2022).

[219] Katharina Ribbeck and Dirk Görlich. “Kinetic Analysis of Translocation through Nuclear Pore Complexes”. In: *The EMBO Journal* 20.6 (2001), pp. 1320–1330. DOI: 10.1093/emboj/20.6.1320.

[220] Alessandro Ori et al. “Cell Type-specific Nuclear Pores: A Case in Point for Context-dependent Stoichiometry of Molecular Machines”. In: *Molecular Systems Biology* 9.1 (2013), p. 648. DOI: 10.1038/msb.2013.4.

[221] Abril Gijsbers et al. “Mycobacterium Tuberculosis Ferritin: A Suitable Workhorse Protein for Cryo-EM Development”. In: *Acta Crystallographica. Section D, Structural Biology* 77 (Pt 8 2021), pp. 1077–1083. DOI: 10.1107/S2059798321007233.

[222] Qing Wang, Christopher P. Mercogliano, and Jan Löwe. “A Ferritin-Based Label for Cellular Electron Cryotomography”. In: *Structure* 19.2 (2011), pp. 147–154. DOI: 10.1016/j.str.2010.12.002.

[223] Jie Zhu et al. “Protein Assembly by Design”. In: *Chemical Reviews* 121.22 (2021), pp. 13701–13796. DOI: 10.1021/acs.chemrev.1c00308.

[224] Suyeong Han and Yongwon Jung. “Artificial Protein Assemblies with Well-Defined Supramolecular Protein Nanostructures”. In: *Biochemical Society Transactions* 49.6 (2021), pp. 2821–2830. DOI: 10.1042/BST20210808.

[225] Bin Yang et al. “Epi-Illumination SPIM for Volumetric Imaging with High Spatial-Temporal Resolution”. In: *Nature Methods* (2019), p. 1. DOI: 10.1038/s41592-019-0401-3.

[226] J. Christof M. Gebhardt et al. “Single-Molecule Imaging of Transcription Factor Binding to DNA in Live Mammalian Cells”. In: *Nature Methods* 10.5 (2013), pp. 421–426. DOI: 10.1038/nmeth.2411.

[227] Florian Schueder et al. “Multiplexed 3D Super-Resolution Imaging of Whole Cells Using Spinning Disk Confocal Microscopy and DNA-PAINT”. In: *Nature Communications* 8.1 (1 2017), p. 2090. DOI: 10.1038/s41467-017-02028-8.

[228] Alexey N. Butkevich et al. “Triarylmethane Fluorophores Resistant to Oxidative Photobluing”. In: *Journal of the American Chemical Society* 141.2 (2019), pp. 981–989. DOI: 10.1021/jacs.8b11036.

[229] Jonathan B. Grimm et al. “A General Method to Improve Fluorophores Using Deuterated Auxochromes”. In: *JACS Au* (2021). DOI: 10.1021/jacsau.1c00006.

[230] Jiachao Xu et al. “Automated Stoichiometry Analysis of Single-Molecule Fluorescence Imaging Traces via Deep Learning”. In: *Journal of the American Chemical Society* 141.17 (2019), pp. 6976–6985. DOI: 10.1021/jacs.9b00688.

[231] Katja J. Koßmann et al. “A Rationally Designed Connector for Assembly of Protein-Functionalized DNA Nanostructures”. In: *ChemBioChem* 17.12 (2016), pp. 1102–1106. DOI: 10.1002/cbic.201600039.

[232] Carla Miró-Vinyals et al. “HaloTag Engineering for Enhanced Fluorogenicity and Kinetics with a Styrylpyridium Dye”. In: *ChemBioChem* 22.24 (2021), pp. 3398–3401. DOI: 10.1002/cbic.202100424.

[233] Mahipal Ganji et al. “Quantitative Assessment of Labeling Probes for Super-Resolution Microscopy Using Designer DNA Nanostructures”. In: *ChemPhysChem* 22.10 (2021), pp. 911–914. DOI: 10.1002/cphc.202100185.

[234] Gong Zhang et al. “Illuminating Biological Processes through Site-Specific Protein Labeling”. In: *Chemical Society Reviews* 44.11 (2015), pp. 3405–3417. DOI: 10.1039/C4CS00393D.

[235] Daniel de la Torre and Jason W. Chin. “Reprogramming the Genetic Code”. In: *Nature Reviews Genetics* 22.3 (3 2021), pp. 169–184. DOI: 10.1038/s41576-020-00307-7.

[236] Kanokpol Aphicho, Narongyot Kittipanukul, and Chayasith Uttamapinant. “Visualizing the Complexity of Proteins in Living Cells with Genetic Code Expansion”. In: *Current Opinion in Chemical Biology* 66 (2022), p. 102108. DOI: 10.1016/j.cbpa.2021.102108.

[237] Wuyu Mao et al. “A General Strategy to Design Highly Fluorogenic Far-Red and Near-Infrared Tetrazine Bioorthogonal Probes”. In: *Angewandte Chemie International Edition* n/a.n/a (2020). DOI: 10.1002/anie.202011544.

[238] Philipp Werther et al. “Bio-Orthogonal Red and Far-Red Fluorogenic Probes for Wash-Free Live-Cell and Super-resolution Microscopy”. In: *ACS Central Science* 7.9 (2021), pp. 1561–1571. DOI: 10.1021/acscentsci.1c00703.

[239] Gangam Srikanth Kumar et al. “Superfast Tetrazole–BCN Cycloaddition Reaction for Bioorthogonal Protein Labeling on Live Cells”. In: *Journal of the American Chemical Society* (2021). DOI: 10.1021/jacs.1c10354.

[240] Jonathan B. Grimm et al. “General Synthetic Method for Si-Fluoresceins and Si-Rhodamines”. In: *ACS Central Science* 3.9 (2017), pp. 975–985. DOI: 10.1021/acscentsci.7b00247.

[241] Nicolas Lardon et al. “Systematic Tuning of Rhodamine Spirocyclization for Super-resolution Microscopy”. In: *Journal of the American Chemical Society* (2021). DOI: 10.1021/jacs.1c05004.

[242] Fang Huang et al. “Ultra-High Resolution 3D Imaging of Whole Cells”. In: *Cell* 166.4 (2016), pp. 1028–1040. DOI: 10.1016/j.cell.2016.06.016.

[243] Francisco Balzarotti et al. “Nanometer Resolution Imaging and Tracking of Fluorescent Molecules with Minimal Photon Fluxes”. In: *Science* 355.6325 (2017), pp. 606–612. DOI: 10.1126/science.aak9913.

[244] Jelmer Cnossen et al. “Localization Microscopy at Doubled Precision with Patterned Illumination”. In: *Nature Methods* 17.1 (1 2020), pp. 59–63. DOI: 10.1038/s41592-019-0657-7.

[245] Luciano A. Masullo et al. “Pulsed Interleaved MINFLUX”. In: *Nano Letters* (2020). DOI: 10.1021/acs.nanolett.0c04600.

[246] Pierre Jouchet et al. “Nanometric Axial Localization of Single Fluorescent Molecules with Modulated Excitation”. In: *Nature Photonics* (2021), pp. 1–8. DOI: 10.1038/s41566-020-00749-9.

[247] John R. Allen, Stephen T. Ross, and Michael W. Davidson. “Single Molecule Localization Microscopy for Superresolution”. In: *Journal of Optics* 15.9 (2013), p. 094001. DOI: 10.1088/2040-8978/15/9/094001.

[248] Wesley R. Legant et al. “High-Density Three-Dimensional Localization Microscopy across Large Volumes”. In: *Nature Methods* 13.4 (4 2016), pp. 359–365. DOI: 10.1038/nmeth.3797.

[249] J. Vangindertael et al. “An Introduction to Optical Super-Resolution Microscopy for the Adventurous Biologist”. In: *Methods and Applications in Fluorescence* 6.2 (2018), p. 022003. DOI: 10.1088/2050-6120/aaae0c.

[250] Dylan T. Burnette et al. “Bleaching/Blinking Assisted Localization Microscopy for Superresolution Imaging Using Standard Fluorescent Molecules”. In: *Proceedings of the National Academy of Sciences* 108.52 (2011), pp. 21081–21086. DOI: 10.1073/pnas.1117430109.

[251] Sebastian Hauke et al. “Specific Protein Labeling with Caged Fluorophores for Dual-Color Imaging and Super-Resolution Microscopy in Living Cells”. In: *Chemical Science* 8.1 (2017), pp. 559–566. DOI: 10.1039/C6SC02088G.

[252] Philipp Werther et al. “Live-Cell Localization Microscopy with a Fluorogenic and Self-Blinking Tetrazine Probe”. In: *Angewandte Chemie International Edition* 59.2 (2020), pp. 804–810. DOI: 10.1002/anie.201906806.

[253] Florian Salopiata. *Investigation of Molecular Mechanisms of Therapy Resistance in Lung Cancer by Mathematical Modeling*. Dissertation, Ruperto-Carola University of Heidelberg, 2018. DOI: 10.11588/heidok.00024456.

[254] Alexander Laskin, Vadim Laskin, and Aleksei Ostrun. “Refractive Beam Shapers for Optical Systems of Lasers”. In: *Components and Packaging for Laser Systems*. Components and Packaging for Laser Systems. Vol. 9346. International Society for Optics and Photonics, 2015, 93460R. DOI: 10.1117/12.2080864.

[255] Jay P. Morgenstern and Hartmut Land. “Advanced Mammalian Gene Transfer: High Titre Retroviral Vectors with Multiple Drug Selection Markers and a Complementary Helper-Free Packaging Cell Line”. In: *Nucleic Acids Research* 18.12 (1990), pp. 3587–3596. DOI: 10.1093/nar/18.12.3587.

[256] Johannes Schindelin et al. “Fiji: An Open-Source Platform for Biological-Image Analysis”. In: *Nature Methods* 9.7 (2012), pp. 676–682. DOI: 10.1038/nmeth.2019.

[257] Brendan P. Cormack, Raphael H. Valdivia, and Stanley Falkow. “FACS-Optimized Mutants of the Green Fluorescent Protein (GFP)”. In: *Gene*. Flourescent Proteins and Applications 173.1 (1996), pp. 33–38. DOI: 10.1016/0378-1119(95)00685-0.

[258] Nathan C. Shaner et al. “Improved Monomeric Red, Orange and Yellow Fluorescent Proteins Derived from Discosoma Sp. Red Fluorescent Protein”. In: *Nature Biotechnology* 22.12 (2004), pp. 1567–1572. DOI: 10.1038/nbt1037.

[259] Antje Keppler et al. “Labeling of Fusion Proteins with Synthetic Fluorophores in Live Cells”. In: *Proceedings of the National Academy of Sciences* 101.27 (2004), pp. 9955–9959. DOI: 10.1073/pnas.0401923101.

[260] Antje Keppler et al. “Fluorophores for Live Cell Imaging of AGT Fusion Proteins across the Visible Spectrum”. In: *BioTechniques* 41.2 (2006), pp. 167–175. DOI: 10.2144/000112216.

Appendix A

LABELING EFFICIENCY

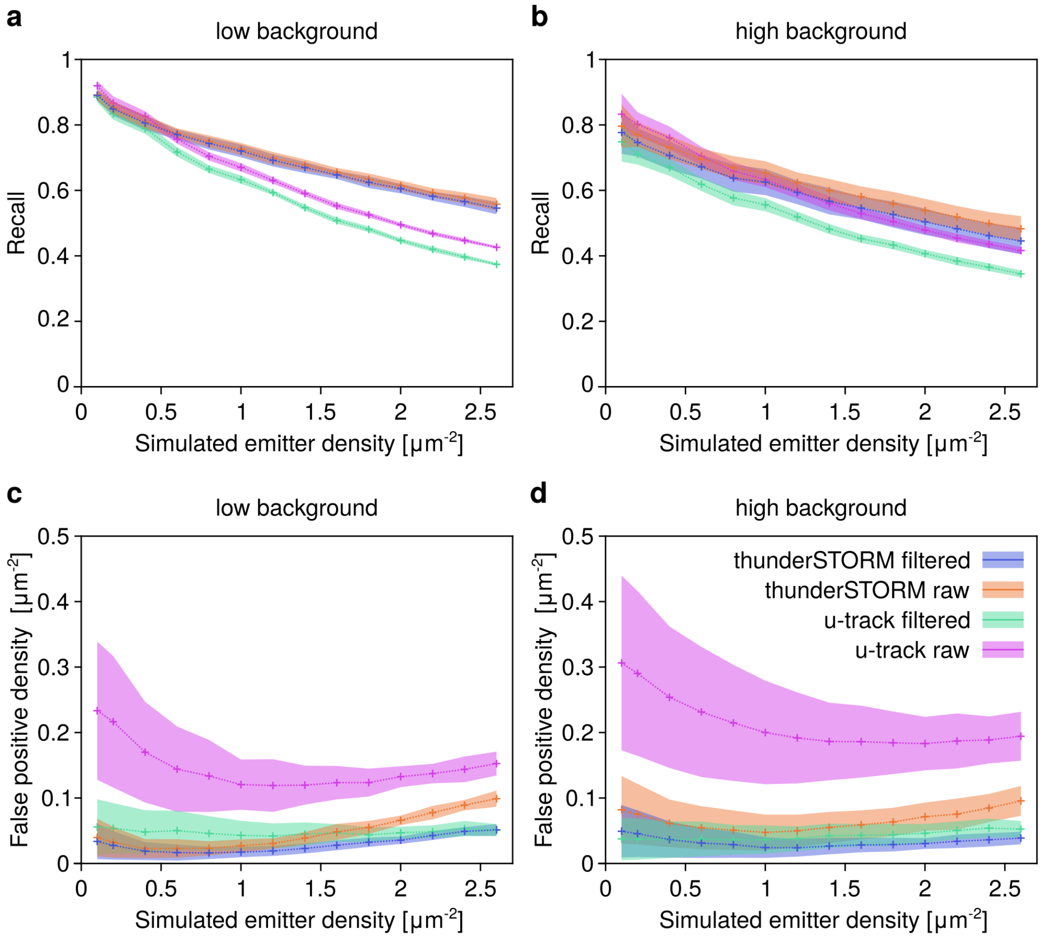

Figure A.1 Recall and false-positive density for tested algorithms with or without localization filtering determined in simulations with variable background strength. **a,b** Recall at variable emitter densities at low (a) and high (b) background signal. **c,d** False-positive density at variable emitter densities at low (c) and high (d) background signal. Shaded regions indicate standard error around means (crosses) at indicated emitter densities. For each algorithm, 20 simulated images per condition were analyzed.

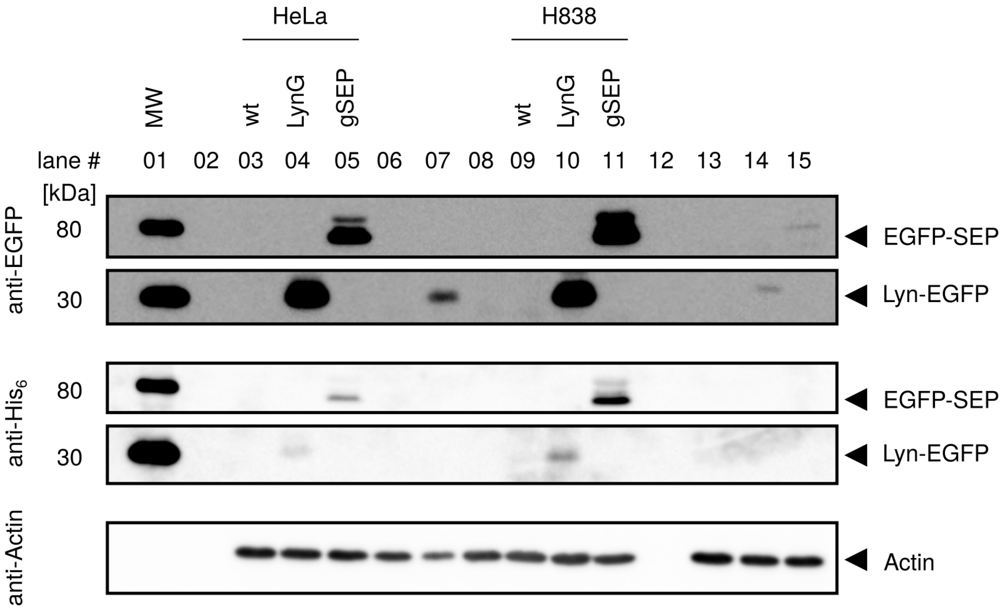

Figure A.2 Full-length expression of gSEP and LynG in H838 and HeLa cell lines confirmed via immunoblotting. H838 or HeLa cells stably selected with puromycin ($1.5\,\mu g\,L^{-1}$ were lysed and subjected to SDS PAGE and and transferred onto a PVDF membrane. The membrane was consecutively probed with anti-EGFP and anti-His$_6$ to verify full-length expression of gSEP and LynG. Probing with anti-actin served as loading control. Lane occupation as indicated above respective lane. Lane 1 was loaded with a molecular size standard, lanes 6-8 and 13-15 contained samples not relevant for this study. Lanes 2 and 12 were not loaded with sample. Data was produced jointly with Florian Salopiata.

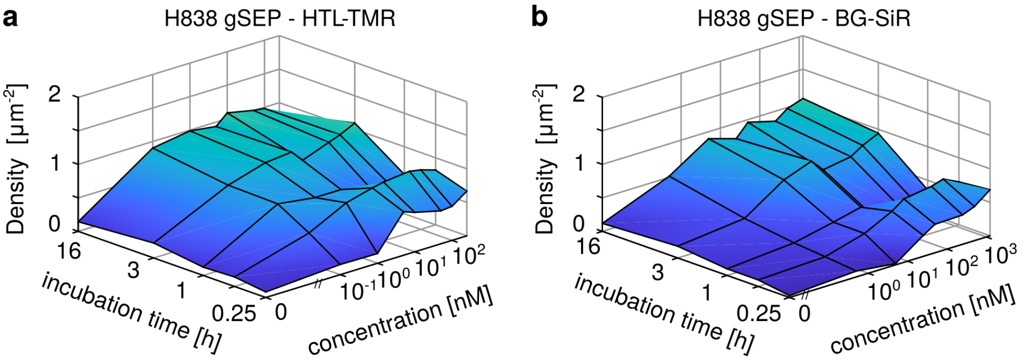

Figure A.3 Target channel emitter densities in H838 gSEP cells. Detected median emitter densities after localization filtering with HTL-TMR (a) or BG-SiR (b). Data from 4-15 cells per condition.

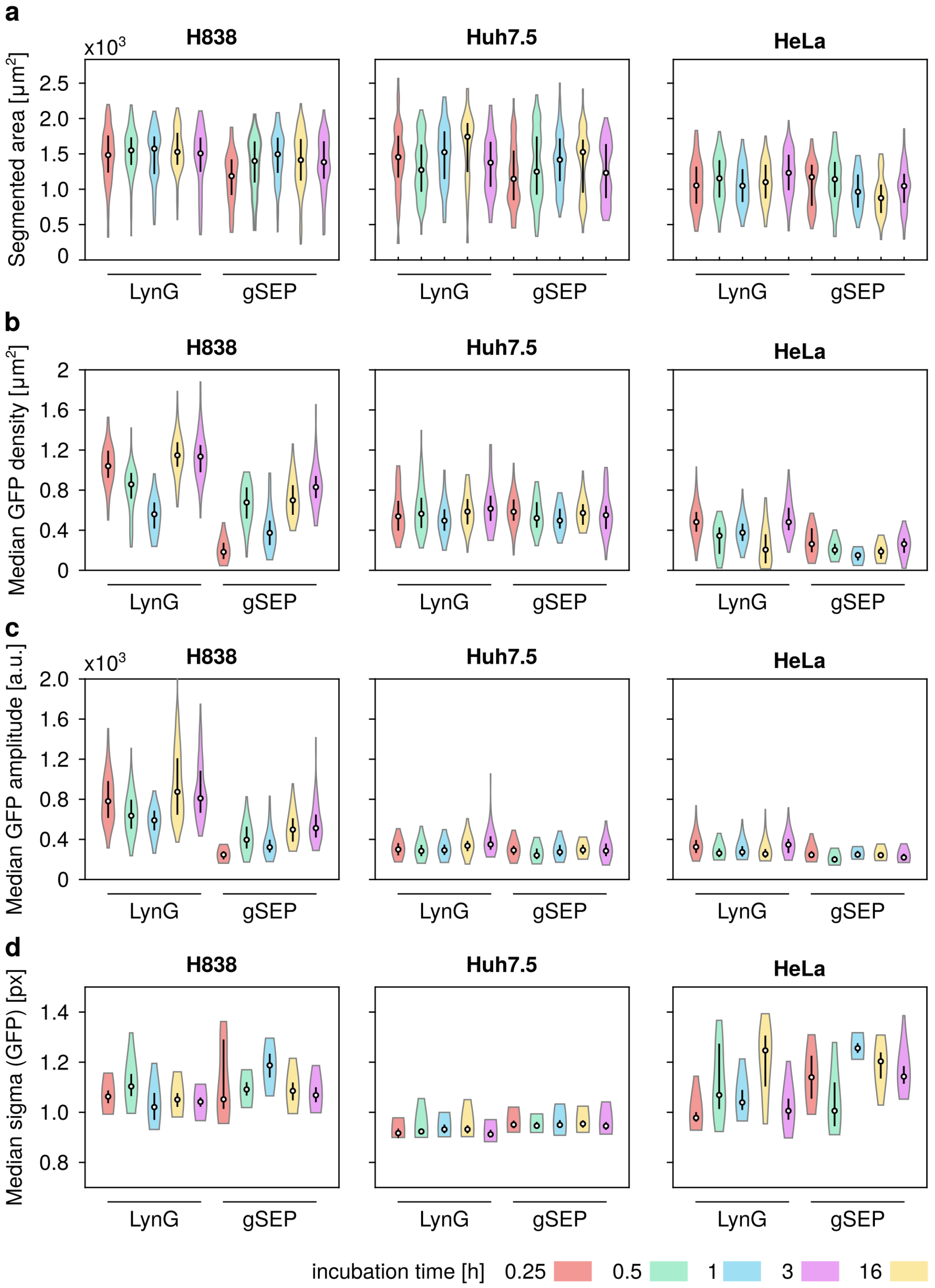

Figure A.4 Quality metrices for cell segmentation and point detection during DOL measurements across cell lines. Metrices were determined for each imaging run which was one 8-well chambered coverslip containing samples incubated with eight different substrate concentrations at a single incubation time. Samples with H838 and HeLa cells were acquired using a semi-automated acquisition routine, samples with Huh7.5 cells were imaged manually. **a,** Cell-wise segmented area. The entire FOV imaged had an area of 2 300-2 800 µm^2 indicating cells typically covered 30-60 % of the FOV. **b-d** density per cell (b), median cell-wise amplitude (c) and median cell-wise sigma (d) of detected points in reference (EGFP) channel. The Huh7.5 dataset was recorded by Siegfried Haenselmann, the HeLa dataset was recorded by Wioleta Chmielewicz. All data was by me.

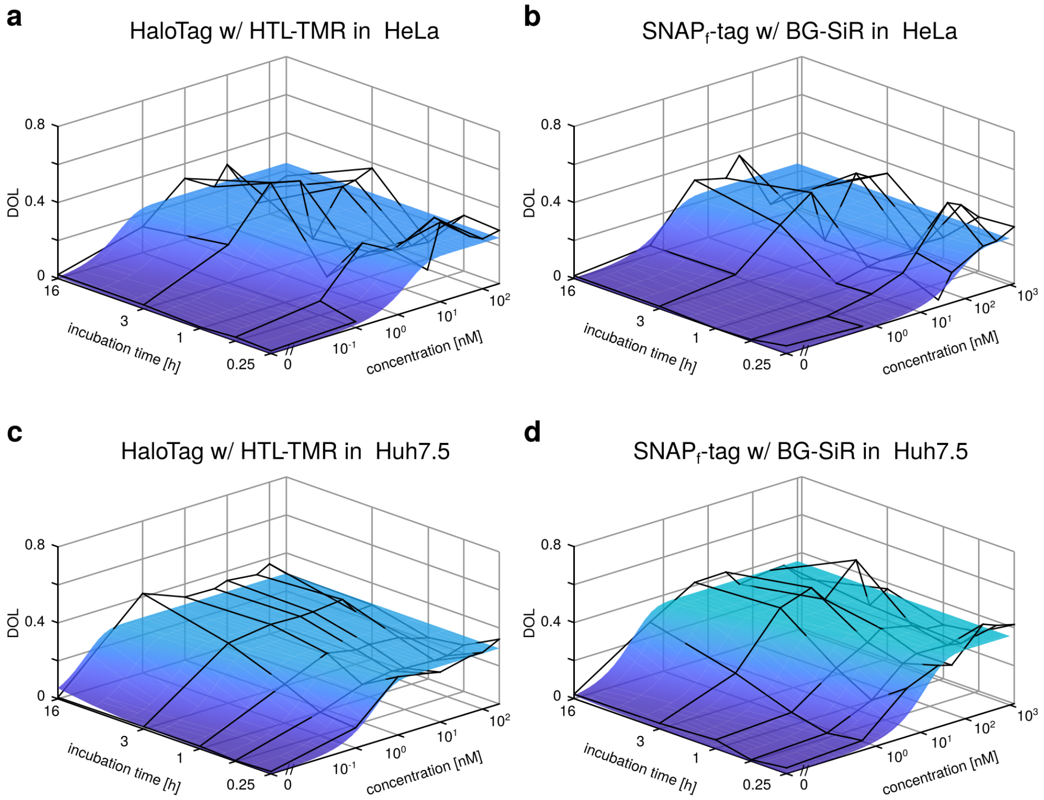

Figure A.5 Comparison of measured DOLs and predicted DOLs from fit with pseudo-first order model (Equation 3.5). Median measured DOLs in indicated gSEP cells (mesh) and predicted DOL (colored surface) obtained from fit for SNAP$_f$-tag or HaloTag labeling with BG-SiR or HTL-TMR. 1-20 (HeLa) and 5-10 (Huh7.5) cells per condition in HeLa and Huh7.5. The Huh7.5 dataset was recorded by Siegfried Haenselmann, the HeLa dataset was recorded by Wioleta Chmielewicz. All data was processed and analyzed by me.

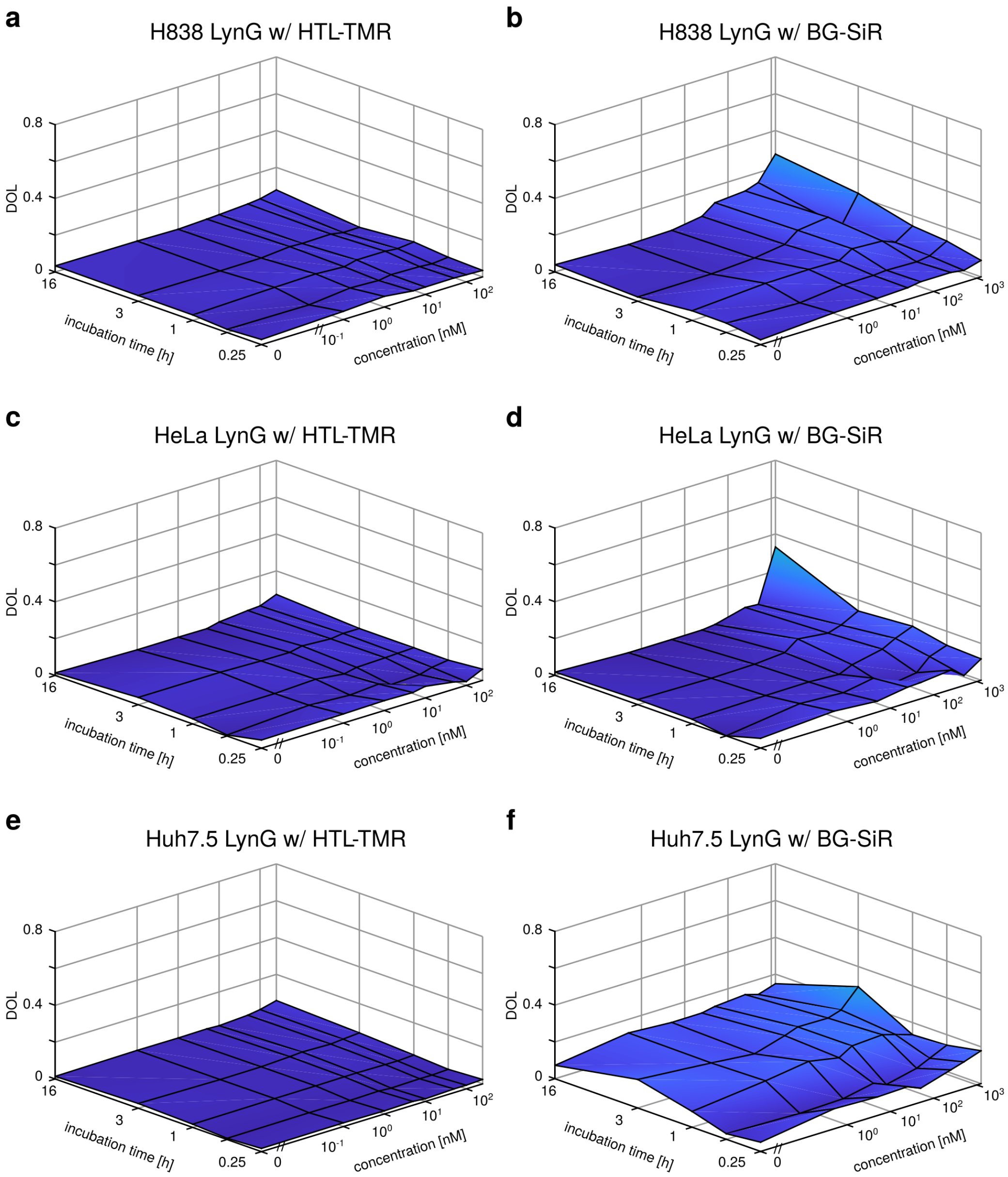

Figure A.6 Apparent DOLs due to non-specific background labeling. Apparent labeling efficiencies due to random co-localization of non-specifically attached tag substrates with LynG probes measured in indicated cell lines with indicated substrates. 3-15 (H838), 5-20 (HeLa) or 5-10 (Huh7.5) cells per condition. The Huh7.5 dataset was recorded by Siegfried Haenselmann, the HeLa dataset was recorded by Wioleta Chmielewicz. All data was processed and analyzed by me.

Appendix B

IN SITU PROTEIN COUNTING

Schwarz Information Criterion for step detection

The step detection criterion developed by Kalafut and Visscher based on the Schwarz Information Criterion (SIC) is defined as:

$$SIC(j_1,...,j_k) = (k+2)log(n) + n\ln\hat{\sigma}^2_{j_1,...,j_k} + n\ln 2\pi + n \tag{B.1}$$

With step positions $j_1,...,j_k$, n data points within trace and the maximum likelihood estimator of the variance $\hat{\sigma}^2_{j_1,...,j_k}$ defined as:

$$\hat{\sigma}^2_{j_1,...,j_k} = \frac{1}{n}\sum_{i=1}^{k}\sum_{l=j_i}^{j_{i+1}-1}(x_l-\mu_i)^2. \tag{B.2}$$

Posterior for model refinement

Posterior was taken from Tsekouras et al.[123]. All parameters are described in Table B.1 below.

$$\begin{aligned}-2\ln P(\theta|D) = &\sum_{\phi=0}^{K}\left(n_\phi\ln(i\sigma_f^2+\sigma_b^2)+\sum_{l=1}^{n_\phi}\frac{(x_l-i\mu_f-\mu_{bg})^2}{i\sigma_f^2+\sigma_b^2}\right)\\ &+2\left(-K\ln\lambda-\ln((m-K)!)-\ln K!-\ln(m-1)!+\sum_{y=0}^{m-K}\ln d_y!\right)\\ &+2\left(\gamma_0\frac{m-K+1}{K}+\ln(m-K+2)+\ln(m-K+1)\right.\\ &\left.-\ln(m-K+2)-(m-K+1)e^{-\gamma_0/K})\right)\end{aligned} \tag{B.3}$$

Table B.1 Symbols used for computation of posterior according to Equation B.3.

θ	Bayesian parameters
D	Data (photobleaching trace)
σ_f	Standard deviation of the single fluorophore signal
σ_{bg}	Standard deviation of the background signal
μ_f	Mean signal of a single fluorophore
μ_{bg}	Mean background signal
K	Number of steps
n_ϕ	Time points in interval ϕ
x	Signal at time point l
i	Active number of fluorophores at timepoint
λ	Poisson distribution parameter for event occurrences (fixed to 0.1)
m	number of events (single fluorophore switches)
d_y	number of data points between steps
γ_0	Cutoff for the hyperparameter constraining K to m (fixed to 0.5)

Table B.2 Overview of parameters for quickPBSA photobleaching step detection framework. Module 3: Trace extraction. Module 4: Preliminary step detection. Module 5: Trace filtering. Module 6: Model refinement. For parameters used in model refinement module, the step within the algorithm is specified in the beginning of the description (I-IV).

Parameter	Module	Default	Description
r_peak	extraction	-	Radius of ROI for trace extraction [pixel].
r_bg1, r_bg2	3	-	Inner and outer radius of background region [pixel] (see Fig. 4.27).
min_dist	3	-	Minimum distance of ROI centers.
threshold	4,5	-	Minimum single-fluorophore brightness.
maxiter	4	-	Maximum number of steps placed during preliminary step detection.
bgframes	4	500	Number of data points to include after crop-point.
length_laststep	5	20	Minimum length of interval between last two steps.
percentile_step	5	90	Upper boundary of last step height (μ_f).
combcutoff	6	2×10^6	Maximum total number of models to test per trace.
multstep_fraction	6	0.5	I - Fraction of steps allowed to be >1.
nonegatives	6	false	I - Do not test step heights >1 for positive steps (fluorophores returning from transient dark states).
mult_threshold	6	1.0	I,II - Threshold for step positions to be tested for heights >1. For occupancies >2, this threshold is applied proportionally.
maxmult	6	5	II - Maximum height of steps.
maxadded	6	10	IV - Maximum number of steps to add.

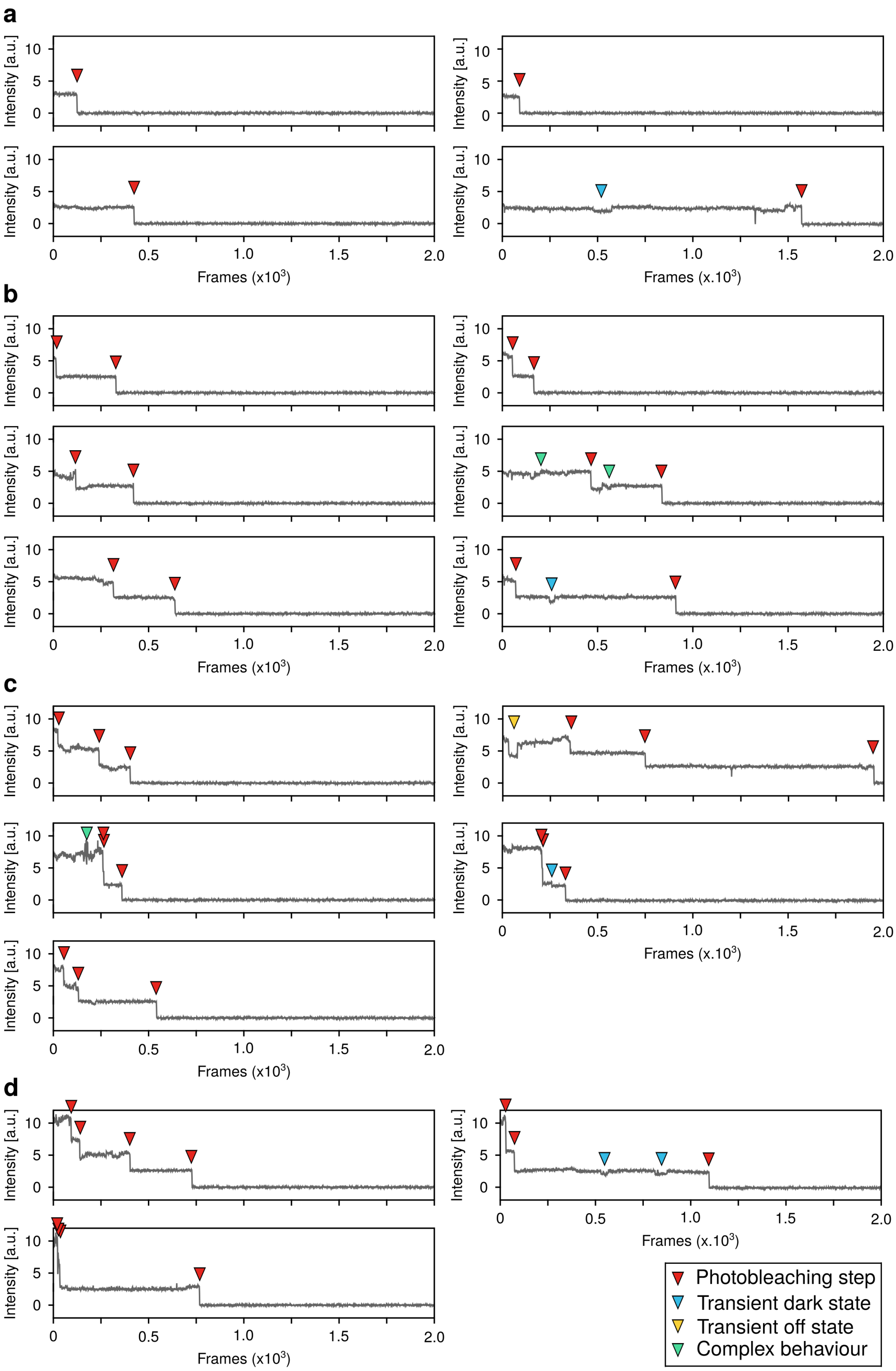

Figure B.1 Classified photobleaching traces used for construction of semi-synthetic traces. Data was recorded from tetraATTO647Nprobes as described in section 4.3. Traces exhibiting different photophysical behavior (arrowheads) were manually categorized according to the number of observed photobleaching steps. Only traces which showed complete bleaching during the measurement were considered. **a,** one photobleaching step, **b,** two photobleaching steps, **c,** three photobleaching steps and **d,** four photobleaching steps.

Table B.3 PBSA results from DNA origami samples. Mean (μ) and SD (σ) for PBSA using indicated algorithms were obtained by Gaussian fitting of fluorophore number estimate distributions. Fitting error given in parentheses. Expected (exp.) mean and SD assuming binomial fluorophore number distribution and a LE of 70 %.

Sample	Setup	Mean (μ)				Standard deviation (μ)				qPBSA success [%]	# traces (FOVs, experiments)	qPBSA runtime [s]
		exp.	qPBSA	KV_{only}	local Int.	exp.	qPBSA	KV_{only}	local Int.			
R09	A	6.3	6.7(0.2)	5.9(0.1)	6.6(0.1)	1.4	1.9(0.2)	2.0(0.1)	2.1(0.1)	75.6	192 (13,1)	25
	B		6.0(0.1)	5.4(0.0)	5.9(0.1)		2.2(0.1)	2.0(0.0)	2.2(0.1)	76.5	1 536 (5,1)	19
R20	A	14.0	14.2(0.3)	11.1(0.3)	15.0(0.6)	2.1	6.0(0.3)	4.3(0.3)	6.4(0.6)	40.0	613 (43,2)	53
R35	A	24.5	22.6(1.0)	16.0(0.7)	23.3(0.8)	2.7	8.6(1.0)	5.8(0.7)	8.9(0.8)	70.5	499 (25,2)	168
Y09	A	6.3	7.9(0.1)	7.5(0.1)	7.9(0.1)	1.4	1.6(0.1)	1.5(0.1)	1.6(0.1)	83.5	853 (12,1)	88

Table B.4 Properties of fluorophores used in Section 4.5.1. Properties determined in PBS, pH 7.4-7.5 unless stated otherwise. εextinction coefficient, Φ_{fl} fluorescence quantum yield, λ_{peak} peak absorption wavelength, λ_{ex} excitation wavelength used in this work, CF_λ spectral correction factor to adjust ε for mismatch in λ_{peak} and λ_{ex}.

	$\varepsilon[10^3\,\mathrm{M}^{-1}\,\mathrm{cm}^{-1}]$	Φ_{fl}	λ_{peak} [nm]	λ_{ex} [nm]	CF_λ	Ref
Fluorescent proteins						
EGFP	55.9	0.6	488	488	1.00	[257]
mNeonGreen	116	0.8	506	488	0.62	[206]
mCherry	72	0.22	587	561	0.64	[258]
Tag substrates						
BG-TMR	$89^{a,e}$	$0.39^{b,f}$	555^{e}	561	0.86^{e}	[259]
HTL-TMR	78	$0.41^{d,e}$	548^{e}	561	0.89^{e}	[74]
BG-SiR	$43.2^{c,f}$	$0.30^{b,f}$	650^{f}	640	0.7^{e}	[77]
HTL-SiR	$130.2^{c,f}$	$0.39^{b,f}$	648^{f}	640	0.71^{e}	[77]
Organic fluorophores						
ATTO 565	120	0.9	564	561	0.99	1
ATTO 647N	150	0.65	646	640	0.92	1

[a] from Keppler et al. (2006)[260]
[b] from Lukinavicius et al. (2013)[77]
[c] from Erdmann et al. (2019)[207]
[d] tetramethylrhodamine, from Grimm et al. (2015)[183]
[e] free dye
[f] protein conjugate
[1] https://www.atto-tec.com/

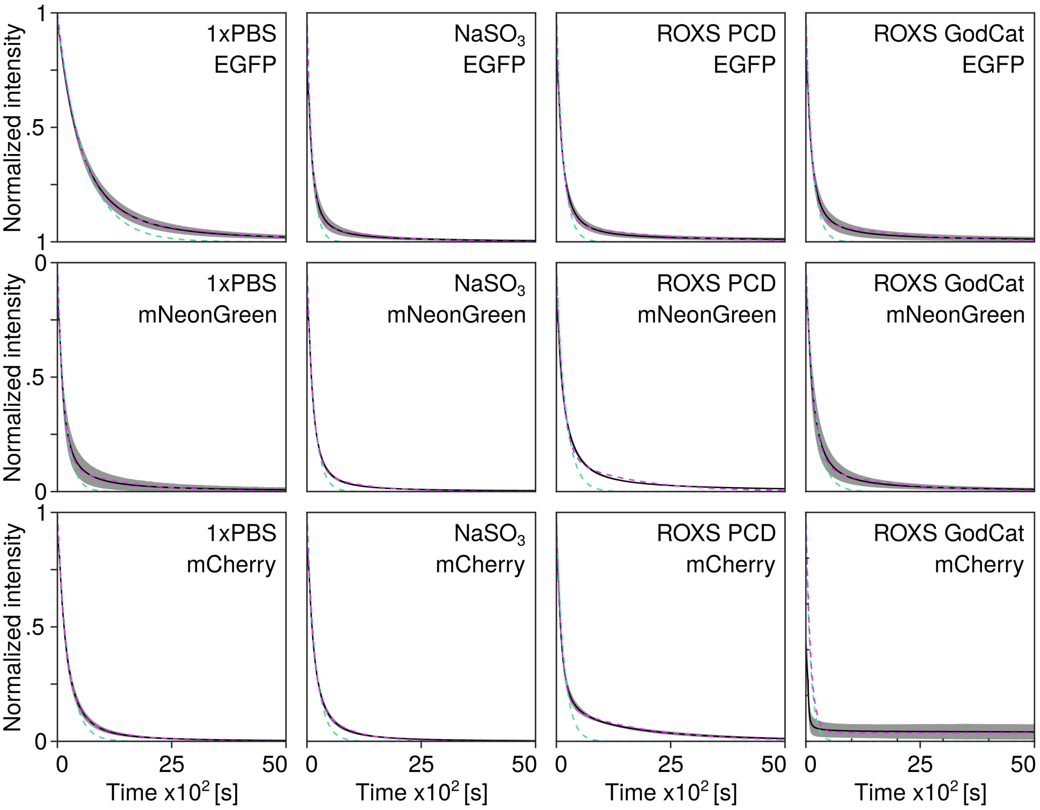

Figure B.2 Photobleaching of fluorescent proteins in ROXS buffers. For each combination of fluorescent protein and buffer, photobleaching traces under constant illumination were recorded in multiple cells. Decays were normalized against intensity in first frame and corrected for intensity offset which depended on sample type and illumination wavelength. Mean±SD (black line, shaded region) from 7-9 cells from two independent experiments for each condition. Illumination power density for EGFP and mNeonGreen: $0.4\,\mathrm{kW\,cm^{-2}}$ at 488 nm, illumination power density for mCherry: $0.52\,\mathrm{kW\,cm^{-2}}$ at 561 nm. Mean decays were modeled with a mono- (green) and bi-exponential (magenta) decay. Results from exponential modeling are listed in Table B.5.

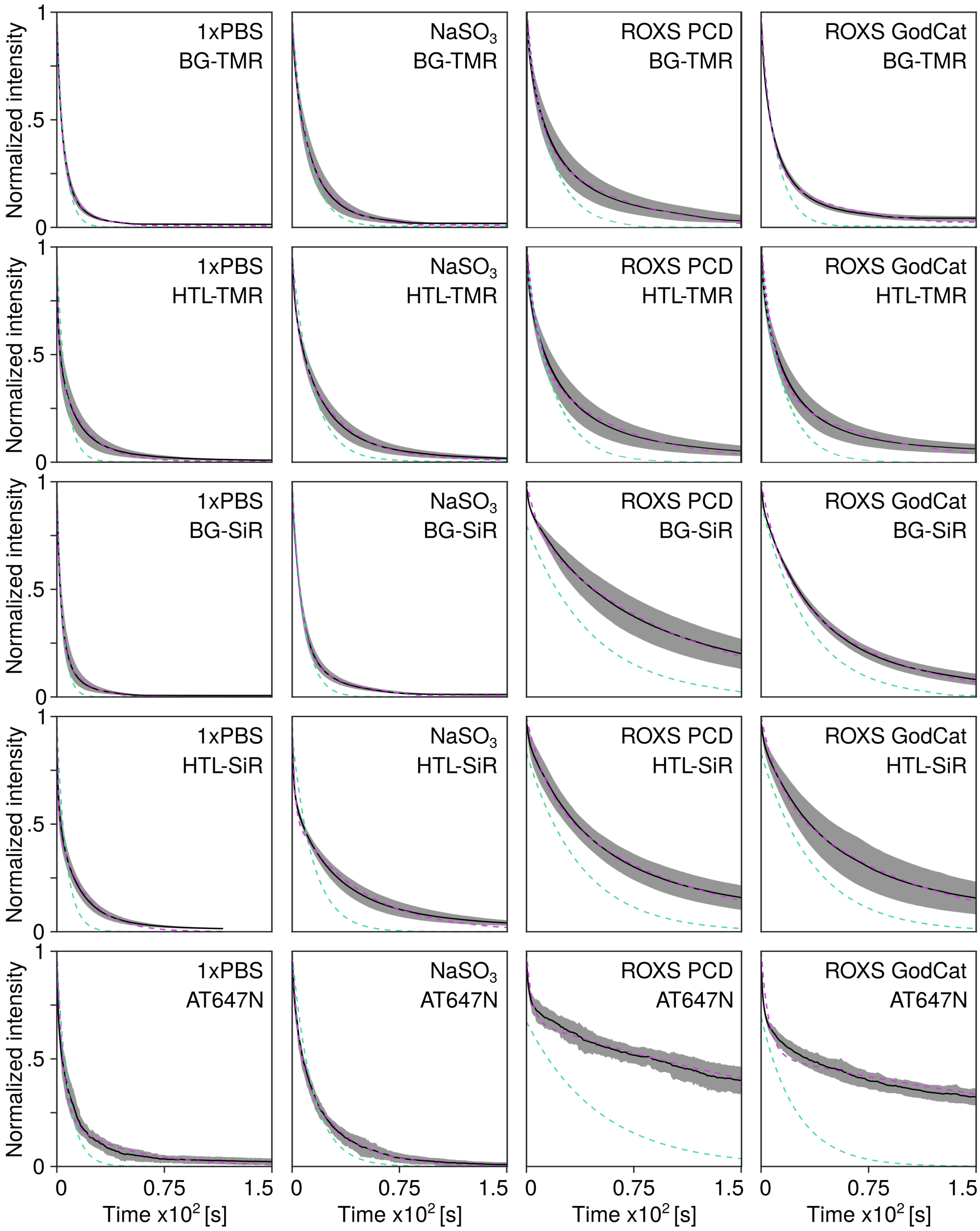

Figure B.3 Photobleaching of organic fluorophores in ROXS buffers. For each combination of dye and buffer, photobleaching traces under constant illumination were recorded in multiple cells (HTL/BG-TMR, HTL/BG-SiR) or fields of view (ATTO647N). Decays were normalized against intensity in first frame and corrected for intensity offset which depended on sample type and illumination wavelength. Mean±SD (black line, shaded region) from 7-9 cells or FOVs from two independent experiments for each condition. Illumination power density for HTL/BG-TMR: 0.52 kW cm^{-2} at 561 nm, illumination power density for HTL/BG-SiR and ATTO647N: 1.26 kW cm^{-2} at 640 nm. Mean decays were modeled with a mono- (green) and bi-exponential (magenta) decay. Results from exponential modeling are listed in Table B.5.

Table B.5 Results from bi-exponential fitting of photobleaching decays shown in Fig. B.2 and Fig. B.3. Decay parameters F1 (fraction fast decay), t1 (fast decay rate), t2 (slow decay rate).

		ROXS		
	PBS	$NaSO_3$	PCD	GodCat
EGFP				
F1	0.837	0.917	0.912	0.904
t1	0.215	1.000	0.816	0.867
t2	0.040	0.075	0.053	0.051
mNeonGreen				
F1	0.881	0.900	0.840	0.886
t1	0.798	0.800	0.725	0.639
t2	0.075	0.105	0.069	0.063
mCherry				
F1	0.855	0.804	0.835	0.963
t1	0.556	0.800	0.962	1.00
t2	0.104	0.147	0.059	0.00
BG-TMR				
F1	0.551	0.414	0.514	0.634
t1	0.765	0.284	0.136	0.115
t2	0.053	0.034	0.017	0.014
HTL-TMR				
F1	0.756	0.455	0.568	0.723
t1	0.287	0.233	0.128	0.156
t2	0.058	0.047	0.018	0.020
BG-SiR				
F1	0.519	0.462	0.337	0.324
t1	0.896	0.583	0.073	0.084
t2	0.051	0.022	0.010	0.010
HTL-SiR				
F1	0.739	0.807	0.216	0.348
t1	0.407	0.168	0.124	0.083
t2	0.070	0.029	0.010	0.015
ATTO647N				
F1	0.767	0.475	0.322	0.481
t1	0.233	0.227	0.272	0.188
t2	0.023	0.035	0.003	0.003

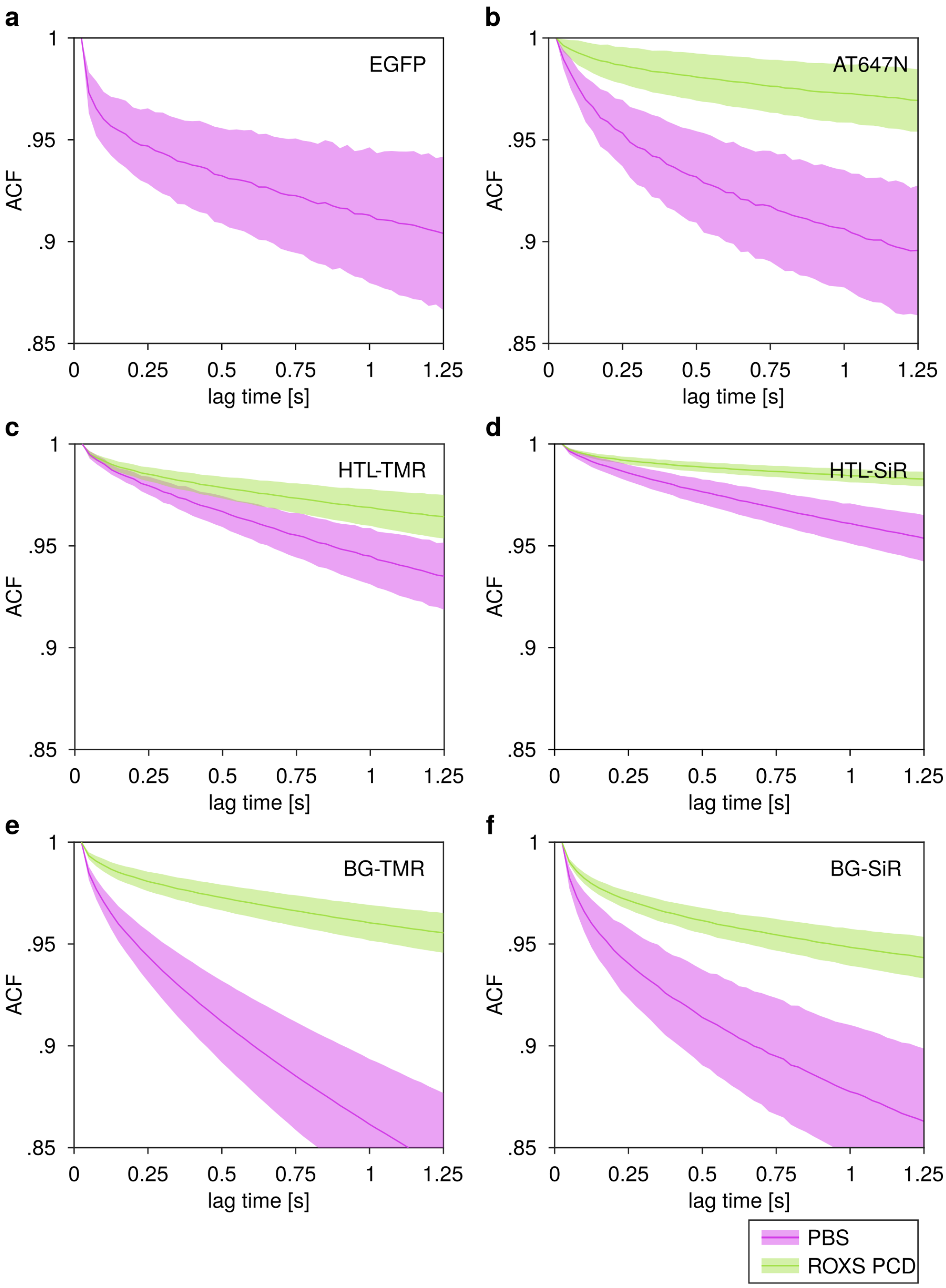

Figure B.4 Temporal autocorrelation functions computed for photoblinking analysis.. **a-e,** Averaged ACFs firstplural (lines) and standard deviation (shaded region) for indicated fluorophores in PBS (magenta) or ROXS PCD (green). Averages from 22-54 measurements from 2-4 independent experiments per condition. Data was produced jointly with Jonas Euchner, Yin Xin Ho and Stanimir Asenov Tashev.

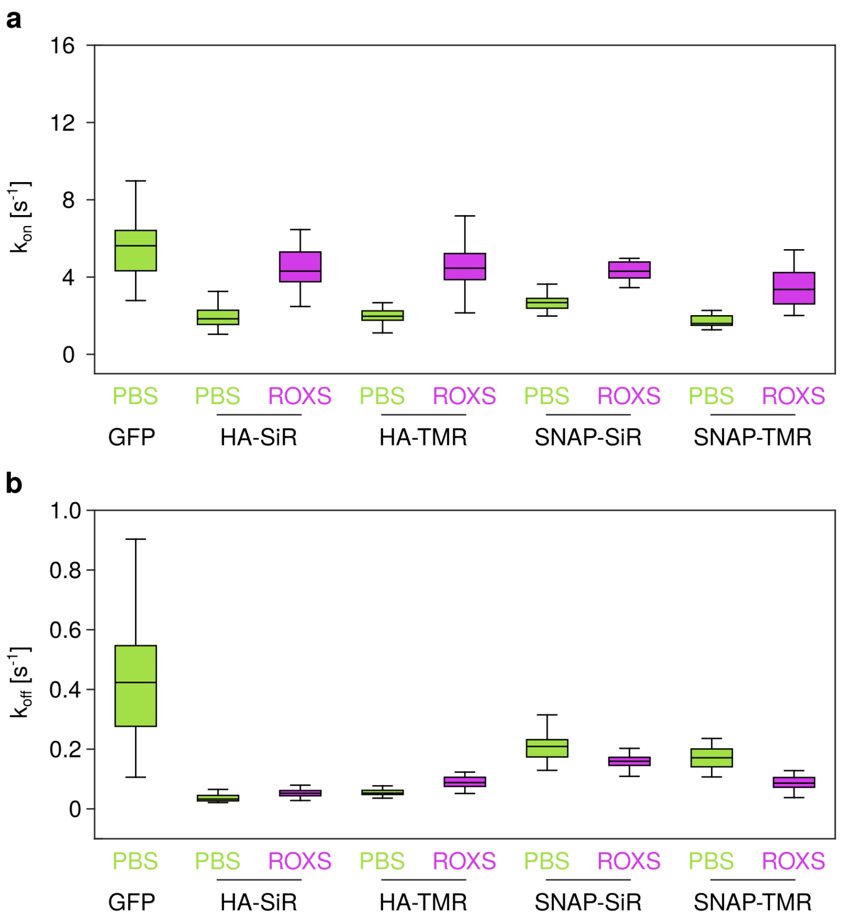

Figure B.5 Rate parameters obtained from image correlation-based analysis of photoblinking. Fluorophore k_{on}- (a) and k_{off}-rates (b) for photoblinking between fluorescent and non-fluorescent state in 3-state fluorophore model obtained from fitting of ACFs firstplural. Data was produced jointly with Jonas Euchner, Yin Xin Ho and Stanimir Asenov Tashev.

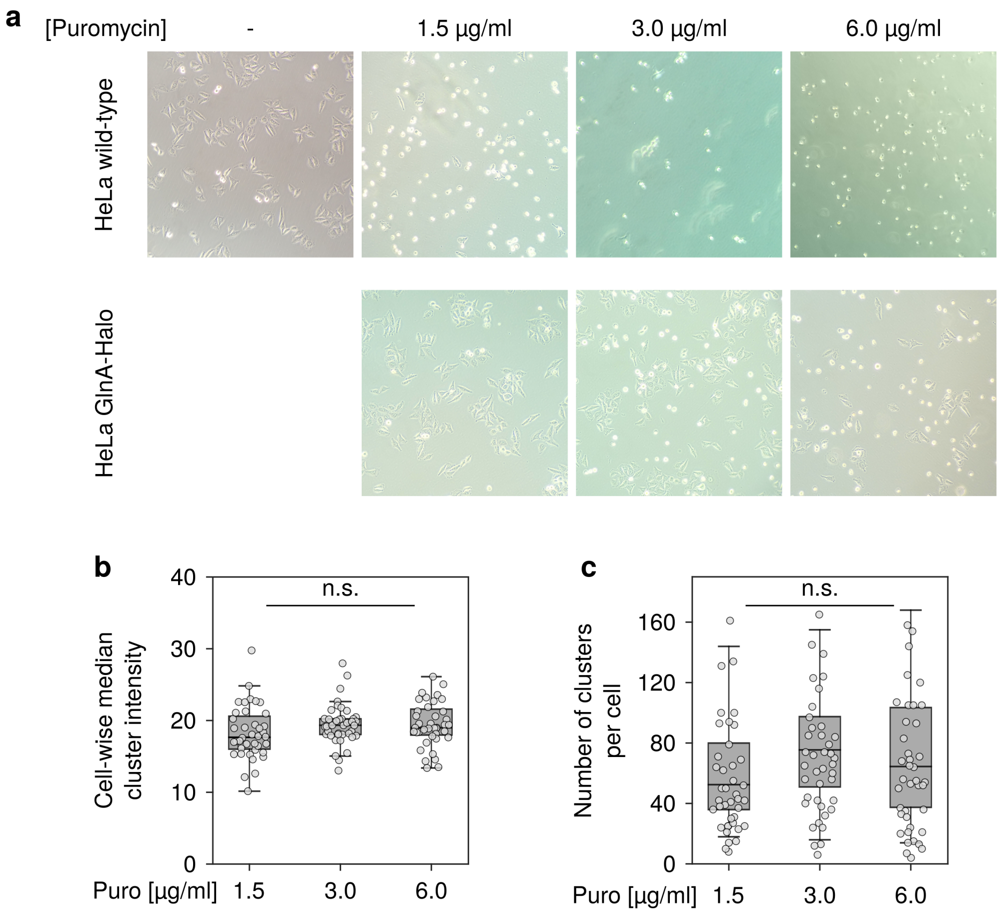

Figure B.6 Puromycin selection of stably transduced HeLa GlnA-HaloTag cell line. **a,** HeLa wild-type and HeLa GlnA-HaloTag cells were grown in full growth medium containing the selection antibiotic at indicated concentrations for 24 hours and imaged with transmission light microscopy. Both cell lines exhibited a puromycin concentration dependent reduction in viability. Two weeks after onset of selection, samples were stained with 10 nm HTL-SiR, fixed in 3.7 % PFA and imaged on epifluorescence setup I. Diffraction-limited clusters were identified by sub-pixel localization and conditions were compared with respect to their median cell-wise cluster intensity (**b**) and cluster number per cell (**c**). Statistical comparison between conditions was performed using Welch's *t*-test. No significant pair-wise differences between conditions were found at $p > 0.05$

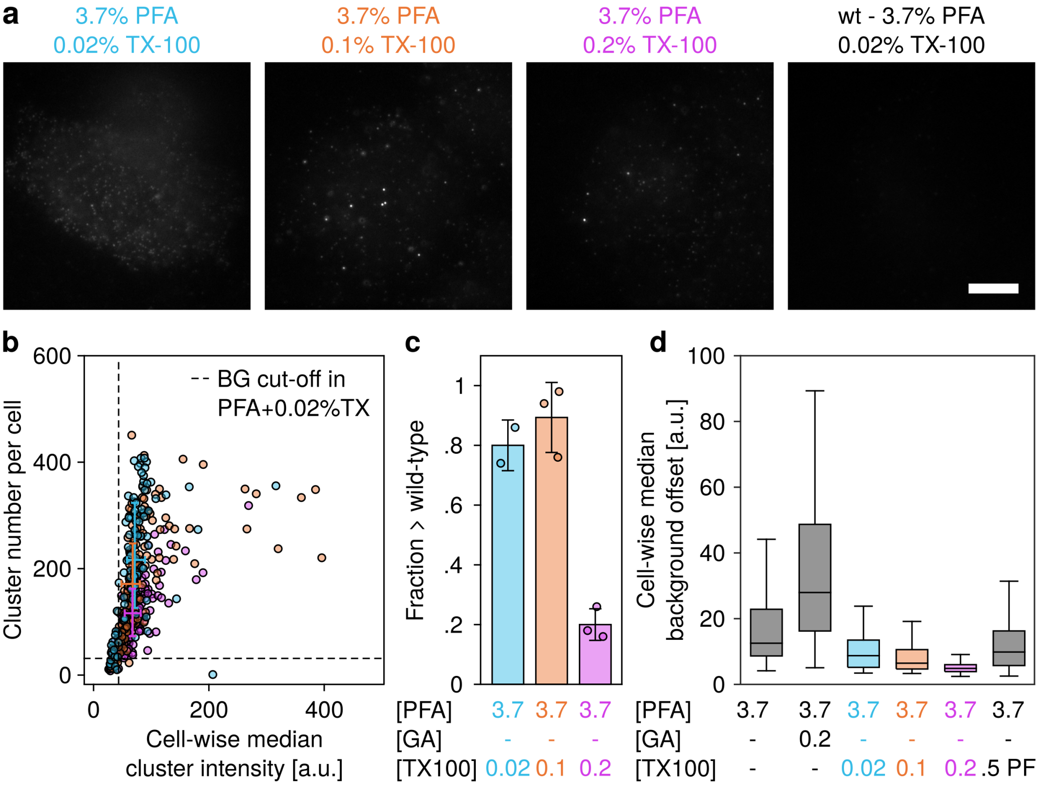

Figure B.7 Influence of Triton X-100 concentration during fixation on GlnA-HaloTag cluster intensity and number. HeLa GlnA-HaloTag cells were labeled with 10 nm HTL-SiR for 60 min, chemically fixed under indicated conditions and imaged by widefield fluorescence microscopy. Cluster properties were obtained by subsequent SMLM analysis. **a,** Representative images from HeLa GlnA-HaloTag cells fixed with 3.7 % PFA supplemented with variable concentrations of TX-100 and HeLa wild-type cells fixed in 3.7 % PFA supplemented with 0.02 % TX-100. Scale bar: 10 µm. **b,** Cell-wise cluster number and median cluster intensity for different TX-100 concentrations (color coding as in a). For each condition, a two-axis error plot indicates mean±SD cluster intensities and cluster number per cell. 95th percentile cut-off for 3.7 % PFA supplemented with 0.02 % TX-100 fixation from wild-type cells indicated by dashed lines (data points not shown in plot). **c,** Fraction of cells in HeLa GlnA-HaloTag samples with median cluster intensity and cluster number >95th percentile of corresponding wild-type sample under identical fixation conditions. Bar represents mean across three samples, error bar represents SD. **d,** Background intensities under varying fixation conditions measured as the offset from Gaussian fitting of individual clusters in SMLM analysis.