

DISSERTATION

submitted to the
Combined Faculties for the Natural Sciences and for
Mathematics
of the
Ruperto-Carola University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

Presented by
M.Sc. Siegfried Häselmann
born in Ettenheim, Germany

Oral examination: September 29, 2017

**Quantitative Microscopy:
Measuring membrane
receptor interactions in live
cells**

Referees: Prof. Dr. Ursula Kummer
Prof. Dr. Dirk-Peter Herten

Für Sari und den Klonier-Flo.

„Science, my lad, is made up of mistakes, but they are mistakes which it is
useful to make, because they lead little by little to the truth.“

Jules Verne

Contents

Abstract	xi
Zusammenfassung	xiii
I. Introduction and Theory	1
1. Introduction	3
1.1. Motivation	3
1.2. Quantitative Single-Molecule Microscopy	3
1.2.1. Fundamentals of Fluorescence	4
1.2.2. Single-Molecule Fluorescence Microscopy	6
1.2.3. The SNAPf-tag and HaloTag System	7
1.2.4. Single-Molecule Tracking	10
1.3. Role and Dynamics of IFN- α Signaling	11
1.3.1. IFN- α Signaling as an Innate Immune Response	11
1.3.2. Dynamics of IFN- α Receptor Dimerization	13
1.4. EGFR and c-Met Signaling in Non-Small Cell Lung Cancer	14
1.4.1. The Role of EGFR and c-Met in Physiology and Patho- physiology	14
1.4.2. EGFR as a Target in Non-Small Cell Lung Cancer Therapy	15
1.5. Scope of this Study	18
II. Experiments and Results	21
2. The Degree of Labeling of SNAPf-tag and HaloTag	23
2.1. Cloning of the GFP-extended Staining Efficiency Probe (gSEP) . .	25
2.2. Creation of Stable Cell Lines	27
2.3. Determination of Förster Resonance Energy Transfer between the Fluorophore Pairs in the gSEP Construct	29
2.3.1. Acceptor Bleaching to Avoid FRET between eGFP and TMR Within the gSEP Construct	32
2.4. Detection of Single Particles	35

2.5.	Image Registration of the eGFP, TMR and SiR Channels	40
2.5.1.	Development of an Linear Image Registration Based on a Fiducial Markers	40
2.5.2.	Consistency of Fiducial Marker Registration Results . . .	43
2.5.3.	Registration of Cell Images	46
2.5.4.	Determination of the Spatial Colocalization Tolerance . .	53
2.5.5.	Prevention of Multiple Assignments	55
2.6.	The Degree of Labeling of SNAPf-tag and HaloTag	57
2.6.1.	Unspecific Labeling of HaloTag TMR ligand and SiR-BG .	62
2.6.2.	Optimal Labeling of HaloTag and SNAPf-tag for Single- Molecule Microscopy	65
2.6.3.	Degree of Maturation of eGFP	67
3.	Two-Color Single-Molecule Tracking of Type I Interferon Receptor	69
3.1.	Generating a Stable Huh7.5-Halo-IFNAR1/SNAPf-IFNAR2 Cell Line	70
3.2.	Two-Color Single-Molecule Tracking of IFNAR1 and IFNAR2 . .	72
3.2.1.	Single-Molecule Tracking and Registration of Spectral Channels	72
3.2.2.	Estimation of a Distance Threshold for Dimer Classification	76
3.2.3.	Quantification of IFNAR1/IFNAR2 heterodimerization . .	78
4.	Ligand-induced Heterodimerization of EGFR and c-Met	85
4.1.	Generating a stable H838-EGFR-SNAPf/c-Met-Halo cell line . . .	86
4.2.	Two-Color Single-Molecule Tracking of EGFR and c-Met	89
4.2.1.	Heterodimerization Reduces Ligand-dependent Endocytosis of c-Met	93
4.2.2.	Dimer Detection Based on a Hidden Markov Model	93
III.	Discussion and Outlook	97
5.	Discussion and Outlook	99
5.1.	The Importance of DOL Determination for Quantitative Microscopy	99
5.2.	Development of a Two-Color Single-Molecule Tracking System .	103
5.2.1.	Comparison of the Obtained IFNAR Dimerization Dynamic with Literature	103
5.3.	Implications of EGFR/c-Met Dimerization for Lung Cancer Therapy	106

IV. Materials and Methods	109
6. Methods and Reagents	111
6.1. Molecular Cloning	111
6.1.1. Polymerase Chain Reaction	111
6.1.2. Restriction Digest	111
6.1.3. DNA Purification	112
6.1.4. Ligation and Transformation	112
6.1.5. MiniPrep	113
6.1.6. MaxiPrep	113
6.2. Cell Culture	113
6.2.1. Retroviral Transduction	113
6.2.2. Splitting, Freezing and Thawing of Cells	115
6.3. Coupling of Siliconrhodamine Benzylguanine	115
6.4. Cleaning of Labteks	116
6.5. Single-Molecule Alternating Laser Excitation (smALEX) Measurements	116
6.6. Bleaching of eGFP and Halo-TMR in fixed Huh7.5-gSEP cells	118
6.7. Determination of the DOL of HaloTag and SNAPf-tag	119
6.7.1. Beads samples	119
6.7.2. Simulation of Single-Molecule Images	119
6.7.3. Cell Staining for the Determination of the DOL of HaloTag and SNAPf-tag	119
6.7.4. Three-Color Single-Molecule Microscopy of fixed cells	120
6.7.5. Analysis of DOL	121
6.8. Single Molecule Tracking of Labeled Receptors	121
6.8.1. Cell Staining for Single-Molecule Tracking	121
6.8.2. Two-Color Live Cell Single-Molecule Tracking	122
6.8.3. Bootstrap Significance Test of Dimer Fraction.	123
6.8.4. Bootstrap Significance Test of Dissociation Rates and Calculation of the Effect Size.	123
6.9. Time-lapse Imaging to Determine Receptor Internalization	124
7. Equipment	125
7.1. Nikon TIRF microscope	125
7.2. Software	125

V. Appendix	127
A. Appendix	129
A.1. Cloning of the gSEP and LynG construct	129
A.2. Dimerization-dependent reduction of particle mobility	132
A.3. Bell-shaped equilibrium binding curves for dimerizing receptor systems	134
A.4. Plasmid Sequences	137
Abbreviations	143
Bibliography	144
Publications	153
Acknowledgments	155
Eidesstattliche Versicherung	157

Abstract

Quantitative fluorescence microscopy is increasingly promoting the understanding of cellular processes on a molecular level. With many of these processes happening on short timescales, low frequencies and in chemical equilibrium, single-molecule techniques provide the necessary resolution and sensitivity to unravel molecular dynamics. To determine protein diffusion as well as protein-protein interactions in live cells, single-molecule tracking is one method of choice.

In this study, I have established a two-color single-molecule tracking system to detect and quantify receptor-receptor interactions in the plasma membrane of live cells based on SNAPf-tag and HaloTag labeling. As a proof-of-function, I could verify the well-described ligand-induced heterodimerization of the type I Interferon receptor and determine its interaction dynamics at physiological conditions. In the clinically relevant setting of lung cancer therapy, I could directly proof the prediction of a mathematical model¹, upon which the direct interaction of epidermal growth factor receptor (EGFR) and hepatocyte growth factor receptor (c-Met) critically tunes cell signaling and sensitivity against tyrosine kinase inhibitors (TKIs) in non-small cell lung cancer (NSCLC) cells. This has led to the proposal of the EGFR/c-Met expression ratio as a relevant biological marker for TKI responsiveness in advanced NSCLC patients.

To benchmark the two-color single-molecule tracking system and facilitate the quantification of microscopy data in general, I have developed a modular artificial protein (gSEP) that could be used as a monomer and dimer control in tracking experiments, as well as to determine the degree of labeling (DOL) of protein tags and fluorescent proteins. Applying 40 different staining conditions, I found that at most 40 % of SNAPf-tags and 50 % of HaloTags could be fluorescently labeled. As the DOL is a crucial, yet hard-to-determine, correction factor for quantitative single-molecule microscopy, the artificial gSEP protein represents a valuable and versatile tool for the quantification of microscopy data within a cellular system and on a single-molecule level.

¹developed in the groups of Ursula Klingmüller (German Cancer Research Center Heidelberg) and Jens Timmer (Freiburg University)

Zusammenfassung

Im Hinblick auf das molekulare Verständnis zellulärer Prozesse hat die quantitative Mikroskopie eine Schlüsselrolle eingenommen. Viele zelluläre Prozesse finden verhältnismäßig selten, auf kurzen Zeitskalen und im chemischen Gleichgewicht statt. Bei der Entschlüsselung dieser Dynamiken bieten Einzelmolekültechniken die nötige Auflösung und Sensitivität. Einzelmolekültracking im Speziellen wurde bereits vielfach eingesetzt, um die Diffusion und Interaktion von Proteinen in lebenden Zellen zu messen.

In dieser Forschungsarbeit habe ich ein zwei-Farben Einzelmolekültracking System basierend auf SNAPf-tag und HaloTag Färbung aufgebaut, um die Interaktion zwischen Rezeptoren in der Plasmamembran lebender Zellen zu quantifizieren. Mit diesem System konnte ich die bereits bekannte und gut beschriebene Heterodimerisierung des Typ I Interferon Rezeptors bestätigen und die Interaktionskinetik des Rezeptorkomplexes unter physiologischen Bedingungen bestimmen. Als klinisch relevante Anwendung auf dem Gebiet der Lungenkrebstherapie konnte ich eine Liganden-abhängige Heterodimerisierung zwischen dem epidermal growth factor receptor (EGFR) und dem hepatocyte growth factor receptor (c-Met) zeigen. Diese Beobachtung bestätigt ein mathematisches Modell¹, nach dem die direkte Interaktion zwischen beiden Rezeptoren sowohl die Stärke des zellulären Wachstumssignals als auch das Ansprechverhalten gegenüber Tyrosinkinase Inhibitoren (TKIs) in nicht-kleinzelligen Lungenkrebs (NSCLC) -Zellen beeinflusst. Daraus konnte das Expressionsverhältnis zwischen EGFR und c-Met als neuer biologischer Marker für das Therapiedesign bei fortgeschrittenem nicht-kleinzelligem Lungenkrebs vorgeschlagen werden.

Um den Detektionsbereich des zwei-Farben Einzelmolekültracking Systems zu bestimmen, habe ich ein künstliches Protein (gSEP) entwickelt, das je nach Färbeprotokoll als Monomer- oder Dimerkontrolle benutzt werden kann. Es stellt außerdem ein vielseitiges Werkzeug dar, um die Färbefizienz von Proteintags und fluoreszenten Proteinen zu bestimmen. Die Färbefizienz ist ein wichtiger, jedoch schwer zugänglicher Korrekturfaktor für die quantitative Mikroskopie. Basierend auf 40 verschiedenen Färbbedingungen konnte ich feststellen, dass

¹entwickelt in den Arbeitsgruppen von Ursula Klingmüller (Deutsches Krebsforschungszentrum Heidelberg) und Jens Timmer (Universität Freiburg)

höchstens 40 % aller SNAPf-tags und 50 % aller HaloTags tatsächlich fluoreszent sichtbar sind.

Part I.

Introduction and Theory

Chapter 1.

Introduction

1.1. Motivation

Understanding a biological system in detail is a requirement to treat or prevent malfunctions. For many signaling pathways in health and disease, cellular signaling by membrane receptors is a major regulatory hub. This is especially true for cancer biology and the immune system, where receptor signaling plays an important role in regulating cell growth, sensing pathogens and facilitating cell-cell communication. The dynamics and interactions of membrane receptors crucially influence the biological outcome and often happen on short timescales and low frequencies. Such fast dynamics can be resolved and quantified in living cells using quantitative single-molecule sensitive techniques. The combination of fluorescence microscopy-based single-molecule tracking with quantitative labels in multiple colors allows to detect and quantify single receptor dimerization events on a millisecond timescale in live cells. In this study, I have established and characterized a two-color single-molecule tracking setup based on HaloTag and SNAPf-tag labeling, in order to shed light on receptor-receptor interactions relevant to Interferon- α signaling and the sensitivity of non-small cell lung cancer cells against tyrosine kinase inhibitors (TKI).

1.2. Quantitative Single-Molecule Microscopy

In the last decades, fluorescence microscopy techniques have rushed into biological research. With the discovery and characterization of the green fluorescent protein (GFP), a protein that gives the hydromedusa *Aequorea victoria* and many other marine organisms a greenish glow [1], by Osamu Shimomura and his co-workers [2, 3], biologist were handed a tool to visualize proteins inside living cells or even in whole organisms. Until now, many other fluorescent proteins with different properties for different demands have been developed [4, 5] and accompanied by many alternative labels including organic dyes or semiconductor quantum dots [6, 7]. Unlike fluorescent proteins, which are genetically fused to a target protein,

these non-genetically encoded fluorophores need to be coupled to their target protein with the help of antibodies, unnatural amino acids [8] or protein tags [9], among others. With an increasing brightness and photostability of fluorescent labels, the sensitivity of fluorescence microscopy has reached a single-molecule level, which allows to count individual molecules or follow their movement in real time. At this ultimate level of molecular resolution, much can be learned about the stochasticity but also general rules of molecular dynamics.

1.2.1. Fundamentals of Fluorescence

Many quantitative microscopy techniques are based on fluorescence microscopy. As such, they detect photons that are emitted by fluorescent molecules upon specific excitation. Such an excitation/emission cycle can be explained by a simplified Jablonski diagram (Figure 1.1) [10] that illustrates possible transitions between the energy states of a molecule. Most fluorophores contain a conjugated π -bond system that is populated by two electrons with opposed spin in the electronic ground state S_0 . One of these electrons can be promoted to a higher molecular orbital upon absorption of a photon with a wavelength that matches the energy of this transition. In this vertical transition, different vibrational states of the excited state S_1 can be reached (Franck-Condon principle). Fast vibrational relaxation to the vibronic ground state of the singlet state precedes the spontaneous electronic relaxation back to the ground state during which a fluorescence photon is emitted. Since some of the initial excitation energy is released by vibrational relaxation or other interactions with neighbouring molecules, e.g. solvent reorganization, the fluorescence photon is generally red-shifted compared to the excitation wavelength. This effect is known as Stokes shift.

Besides emitting a fluorescence photon, there exist several alternative pathways to depopulate the S_1 state. Internal conversion (IC) transfers the molecule from the excited S_1 state to a very high vibrational state of the electronic ground state. The excitation energy is then released as heat during vibrational relaxation. Alternatively, intersystem crossing (ISC), which involves a spin flip of the excited electron, transfers the molecule to a triplet state T_1 . This spin conversion is forbidden however facilitated by spin-orbit coupling. Therefore, a transition to the triplet state is for many fluorophores relatively rare but also long-lived as relaxation from the triplet state to the ground state requires another forbidden spin flip. Due to this relatively long lifetime, radiation-free transitions from the triplet state are favored in liquids. These transitions include the formation of a radical anion or cation state and spin exchange with oxygen. Radiation-free pathways generally facilitate chemical reactions between the fluorophore (radical) and (reactive) oxygen that eventually lead to a destruction of the π -bond system and hence to photobleaching

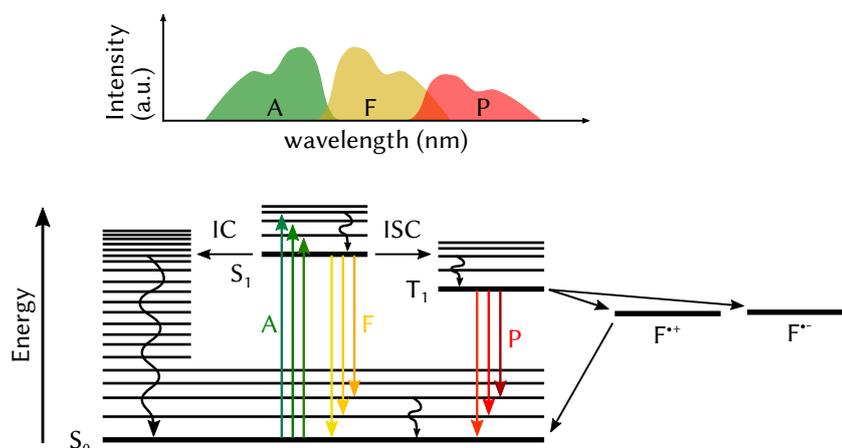


Figure 1.1.: Electronic transitions in a simplified Jablonski diagram. The fluorophore is moved to an electronic excited state (S_1) by absorbing a photon (A). It can return to the electronic ground state (S_0) via internal conversion (IC) or by emitting a red-shifted fluorescence photon (F). Alternatively, it can enter a triplet state via intersystem crossing (ISC) that is depopulated by emitting a phosphorescence photon (P) or forming a radical anion $F^{\bullet-}$ or a radical cation $F^{\bullet+}$.

of the fluorophore. In solids the emission of a phosphorescence photon, which is even further red-shifted, represents a radiative relaxation. Because fluorophores do not emit fluorescent photons while in a triplet state, the triplet state represents a dark or 'off' state and is one reason for fluorophore blinking.

If an excited fluorophore is in close proximity to a different fluorophore in the ground state, their fluorescence dipoles can resonate, which allows the excited donor fluorophore to transfer its energy in a radiation-free process to the acceptor fluorophore [10]. As a result, the donor fluorophore returns to the electronic ground state S_0 , while the acceptor fluorophore is promoted to the excited S_1 state. This process is called Förster resonance energy transfer (FRET) and its efficiency E_{FRET} depends strongly on the distance r between the two fluorophores and their characteristic Förster radius R_0 , which specifies the overlap between the emission spectrum of the donor and the absorption spectrum of the acceptor as well as their spatial orientation (denoted by the orientation factor κ), the quantum yield of the donor fluorophore ϕ_D and the refractive index of the medium n .

$$E_{\text{FRET}} = \frac{R_0^6}{R_0^6 + r^6} \quad (1.1)$$

$$R_0^6 = \frac{9000(\ln 10) \kappa^2 \phi_D}{128\pi^5 N_A n^4} \int f_D(\lambda) \epsilon_A(\lambda) \lambda^4 d\lambda \quad (1.2)$$

While FRET is being used as a very sensitive tool to measure molecular distances [11], it can also present an unwanted quenching mechanisms that reduced the brightness of donor fluorophores, hampering their detection.

1.2.2. Single-Molecule Fluorescence Microscopy

As mentioned, advances in microscopy and labeling techniques in the last decades have increased the spatial resolution and sensitivity to a single-molecule level, an achievement that was honored with the 2014 Nobel price in chemistry [12]. In contrast to classical bulk measurements, which average the signal of many molecules, single-molecule fluorescence microscopy is sensitive to rare and short-lived events and allows to detect heterogeneities within a sample instead of mean values only.

To achieve single-molecule sensitivity, single-molecule fluorescence microscopy requires a selective illumination and a strong background rejection. In confocal microscopes for example, the excitation laser is focused to a relatively small observation volume and a pinhole in the emission pathway selects for light from the focal plane and blocks unfocused background light. For samples where the molecules of interest are located close to the cover slide, total internal reflection fluorescence microscopy (TIRFM) is a robust and widely-used technique to achieve single-molecule sensitivity. Here, the sample is illuminated by the excitation laser in a very shallow angle that leads to a total reflection of the excitation light at the glass-water interface. Due to photon-tunneling however, an evanescent excitation field emerges that allows fluorophores in close proximity to the surface (typically around 100 nm) to be excited. As a consequence, background fluorescence from deeper sample layers is effectively suppressed. Being a wide-field technique, TIRFM allows to image the whole field of view at once (in contrast to confocal microscopy where the field of view has to be sequentially scanned) and thus enables high time resolutions required for live cell tracking. For more complex three-dimensional samples, sophisticated light-sheet fluorescence microscopy-based methods [13] or super-resolution methods (e.g. MINIFLUX [14]) have emerged that provide single-molecule sensitivity and resolution in three dimensions.

Beside its technical requirements, single-molecule fluorescence microscopy also imposes restrictions on the sample and on the fluorescent label in particular. Fluorescent labels should be preferably bright, non-toxic to cells and show a desired photophysics. For single-molecule tracking for example, this would imply a high photostability and no blinking while super-resolution microscopy techniques

require controllable transitions from fluorescent 'on' and 'off' states. Depending on these requirements, single-molecule fluorescence techniques often profit from the use of synthetic organic dyes. Organic dyes generally provide a higher brightness and better photostability compared to fluorescent proteins [9, 15] and they can be chosen from a large set of dyes with different photophysical or chemical properties [16, 17].

Using non-genetically encoded fluorophores however comes at the cost of an additional staining step and the need for a labeling strategy that specifically links the fluorescent dye to the target protein. This task is most commonly taken on by specific antibodies that are decorated with the respective fluorophores. In respect to quantitative microscopy, one disadvantage of using fluorophore-conjugated antibodies is their relatively broad label number, i.e. the number of fluorophores per antibody [18, 19]. This average number of dye molecules per target protein molecule is also referred to as the degree of labeling (DOL). For quantitative microscopy, the degree of labeling is a crucial information as it relates the number of detected fluorophores to the underlying number of target proteins. In order to obtain quantitative information like protein numbers or protein concentrations, it is essential to know how many fluorophores are attached to each target protein because it is the fluorophores that are observed and not the target proteins directly. Therefore, the stoichiometry between the two has to be determined in order to obtain true quantitative information.

In the case of antibody staining, not only the broadly distributed DOL of each antibody but also the varying number of antibodies per target protein turn quantification of single-molecule data challenging. This effect is amplified when a combination of primary antibody and labeled secondary antibody is used for labeling. As will be further described in the following section, protein tags represent an alternative labeling strategy that largely reduces this stoichiometric uncertainty. While multiple antibodies with multiple fluorophores each can bind one target protein, protein tags are genetically fused to the target protein in a 1:1 ratio and they can bind only one fluorophore at most. Therefore they provide a stoichiometric label of the target protein, which facilitates the quantification of single-molecule data.

1.2.3. The SNAPf-tag and HaloTag System

SNAP-tag and HaloTag are the most popular members of a family of protein-based fusion tags. Protein-based fusion tags are usually derived from natural enzymes and designed to bind an organic dye substrate with high affinity or covalently.

SNAP-tag is a 20 kD protein tag that has been derived from the DNA repair enzyme human O6-alkylguanine-DNA alkyltransferase (hAGT) [20]. The endogenous

hAGT recognizes O6-alkylguanine in damaged DNA and catalyzes a dealkylation through a SN_2 suicide reaction mechanism [21]. During this reaction, the enzyme covalently binds the alkyl residue and releases the “repaired”, non-alkylated guanine. For the SNAP-tag system, this reaction is exploited to covalently transfer a fluorophore to the SNAP-tag by using fluorophore-conjugated benzylpurine (BG) or benzylchloropyrimidine (CP) derivatives as a substrate. In the course of the reaction, SNAP-tag covalently binds the fluorophore residue and thus labels the target protein to which it has been genetically fused. SNAP-tag has been derived from hAGT through a process of directed evolution, with the latest version called SNAPf-tag. Within this evolution process, a total of 19 amino acid substitutions and an additional 24-residue deletion at the C-terminus were introduced to increase specificity and speed-up the reaction kinetics more than 10-fold [22, 23]. However, off-target effects of SNAP-tag substrates have also been reported [24].

HaloTag works very similar to SNAP-tag, however it is based on the bacterial 34 kD haloalkane dehalogenase (DhaA) of *Rhodococcus rhodochrous*, which catalyzes the hydrolytic conversion of a haloalkane to the corresponding alcohol and hydrogen halide [25]. As the major difference to the wild type dehalogenase, the active site of HaloTag contains a histidine at position 272 instead of a phenylalanine. This mutation prevents the full release of the substrate as His272 does not support the hydrolysis of a covalent alkyl-enzyme bond that is formed during the reaction [25]. The so created suicide reaction can be exploited to attach a fluorophore to the HaloTag by using a fluorophore-chloroalkane conjugate as a substrate. Through a protein engineering process that utilized site specific and random mutagenesis, the substrate binding kinetics, thermostability and solubility of HaloTag was improved compared to the original DhaA [26]. Since HaloTag is derived from a bacterial enzyme, there is no endogenous equivalent in most biological systems including eukaryotic cells that could compete for HaloTag substrates and generate off-target labeling [26].

Being genetically encoded, SNAP-tag and HaloTag labeling is generally live cell compatible. However, if intracellular proteins are to be labeled, HaloTag and SNAP-tag are restricted to a handful of cell-permeable dye substrates [27]. For fixed cells, these restrictions are lifted as both HaloTag and SNAP-tag can be labeled after fixation and permeabilization of the expressing cell. As HaloTag and SNAP-tag recognize different substrates, they can be used as orthogonal labels.

As mentioned above, SNAP-tag and HaloTag help to reduce the stoichiometric uncertainty between the number of detected fluorophores and the number of underlying target proteins. Due to the enzymatic labeling reaction of the protein tags, they can be labeled with only one fluorophore at most. The genetic fusion between the protein tag and the target protein further insures that only one protein tag is attached to each target protein. In total, this system ensures that each target

protein carries one fluorophore at most. However, because enzymatic reactions as well as protein folding are not 100 % efficient, it can be expected that a certain fraction of protein tags are non-functional or not labeled. This lowers the expected DOL to a value below 1 and requires a thorough quantification of the DOL for any given labeling condition, if absolute protein numbers or protein concentrations are to be determined from microscopy data.

In bulk *in vitro* experiments, the DOL is typically determined by absorption spectrometry as the ratio between the concentration of the dye and the concentration of the target protein, assuming that all dye is bound to the target protein. For SNAP-tag, such experiments have reported a DOL of 0.6–0.7 [28]. Mass spectrometry analysis of *in vitro* labeled SNAP-tag even reported DOLs near 1 [23].

On a single molecule level, the DOL can be determined by fusing an additional and independent reporter to the protein tag. This reporter can be used to localize protein tags independent of their labeling. Using these as a statistical sample of the population of all tags, the DOL can be calculated as the respective proportion of labeled tags (see also chapter 2 on page 23). By calculating the colocalization between *in vitro* labeled SNAP-tags and HaloTags in a fusion construct of both tags, Roder et al. measured a DOL of SNAP-tag and HaloTag of 0.4 and 0.3 on a single-molecule level [29].

In vitro labeling conditions may not necessarily reflect the labeling conditions found in the microenvironment of living cells. As this microenvironment can influence the labeling kinetics of a protein tag as well as the photophysical properties of the dye substrates, it is desirable to determine the DOL of SNAP-tag and HaloTag within the cellular system in which it is used. Combining the idea of cross-referencing the two protein tags with live-cell single-molecule tracking, Wilmes et al. confirmed a DOL of 0.4 and 0.3 for SNAP-tag and HaloTag, respectively [30]. Instead of co-referencing SNAP-tag and HaloTag against each other, Latty and co-workers used an additional antibody labeling to introduce an independent reference signal [31]. In this way, they could estimate a DOL for HaloTag of 0.33 and a DOL of SNAP-tag of 0.16.

These results suggest that SNAP-tag and HaloTag can be robustly labeled but their DOL may be limited to a value well below 1 for *in vivo* labeling. This supports the need for an individual determination of the DOL of protein tags for quantitative microscopy. In addition, a wider screen of *in vivo* labeling conditions would be desirable to better describe the labeling properties of SNAP-tag and HaloTag within a biological system.

1.2.4. Single-Molecule Tracking

Single-molecule tracking (also called single-particle tracking (SPT)) has emerged as a valuable tool to observe the dynamics of proteins *in vitro* and *in vivo* [32]. It usually applies single-molecule sensitive fluorescence microscopy techniques to follow the movement of one or multiple species of molecules in real time. Typically, one-color tracking is being used to answer questions about the motion of molecules (e.g. the transport of virus particles [33]) or about the environment in which molecules move (e.g. the compartmentalization of the cell membrane [34]). Two-color tracking allows to follow two different species of molecules simultaneously and can be used to quantify specific interactions between two proteins in real time [35]. Single-molecule tracking reports on every molecule and is thus sensitive to heterogeneity and rare or short-lived events. This is especially useful for studying transient protein-protein interactions that can hardly be detected using classical ensemble-based methods like co-immunoprecipitation.

Technically, single-molecule tracking requires a single-molecule sensitive fluorescence microscopy technique (e.g. TIRFM for plasma membrane associated proteins) and a specific labeling of the target proteins with bright and photostable fluorophores that show little to no photoblinking. The movement of the labeled target proteins is then imaged with a high frame rate (typically at least 20 Hz).

The analysis of these movies includes three major steps. First, particles have to be detected in every frame of the movie. This is mostly achieved using an intensity-based shape detection (e.g. a two-dimensional Gaussian fit with background comparison). Next, particles in different frames have to be linked in order to generate trajectories. Here, the major parameter to judge if particles of different frames correspond to each other is the distance between them. If a particle is found at the same position in two consecutive frames, it is very likely the same particle and will be linked. Consequently, the quality of the resulting trajectories is generally increased when particles move slowly and at a low density. The introduction of gaps, i.e. frames where a particle is lost but recovered the following frames, can help to close broken trajectories and compensate for fluorophore blinking, for example [36]. A final analysis step is needed to extract information about the motion or interaction dynamics of the particles. Motion analysis can include calculating the diffusion coefficient, the type of motion or the area of confinement, for example. Due to the increasing popularity of one-color single-molecule tracking, there exist several open-source algorithms that can detect particles, create trajectories and perform motion analysis. Arguably the most popular algorithms include the ImageJ plugin TrackMate [37] or the Matlab-based tracking software u-track [36]. Both algorithms scored well in a 2012 Particle Tracking Challenge [38].

For two-color single-molecule tracking, the fluorescent signal of two orthog-

onally labeled protein species is usually spectrally separated and imaged either using two cameras or dedicating half of the camera chip to each spectral channel. In both cases, interaction analysis between the two species requires an image registration step that projects the trajectories of both spectral channels into a common coordinate system. In a next step, interactions between particles can be detected, usually based on the distance between them. The underlying assumption is, that particles that stay in close proximity over time are likely to interact. Such a classification can be implemented with a fixed distance threshold that has to be met for a certain time period [30, 39] or, more sophisticated, by using a hidden Markov model (HMM) with the distance between two particles as the observation layer [35]. Two color single-molecule tracking has already proven useful to determine the interaction kinetics of various cell-surface receptors including EGFR homodimerization [35, 40] or the formation of the type I interferon signaling complex [30].

1.3. Role and Dynamics of IFN- α Signaling

Interferons (IFNs) are pleiotropic cytokines that can be found in all mammalian species and play a role in normal growth and development but more prominently during viral infections and tumor challenge. They are best known for their antiviral properties, with the cellular interferon response being a major defense against intracellular pathogens. Secreted by infected cells, IFNs report on the presence of pathogens and mediate an antiviral and antiproliferative state in neighboring cells (Fig 1.2 (a)). This has turned type I IFNs into a clinical therapy for some viral infections including chronic hepatitis C and hepatitis B infections as well as solid tumors including malignant melanoma and renal cell carcinoma [41].

1.3.1. IFN- α Signaling as an Innate Immune Response

The family of human type I interferons comprises 16 subtypes including IFN- β , IFN- ϵ , IFN- κ , IFN- ω as well as 12 subtypes of IFN- α [42]. The different type I interferons share between 20–60 % sequence homology and all of them bind to the interferon- α receptor (IFNAR) complex [43]. The IFNAR complex belongs to the family of class II helical cytokine receptors and is composed of two individual subchains, interferon- α receptor 1 (IFNAR1) and interferon- α receptor 2 (IFNAR2) [44]. The two receptor chains differ substantially in their binding affinity towards class I interferons. IFNAR1 is generally the low affinity receptor chain with binding affinities in the micromolar range, while IFNAR2 is the high affinity receptor chain with binding affinities in the nanomolar range [42]. As a consequence, it

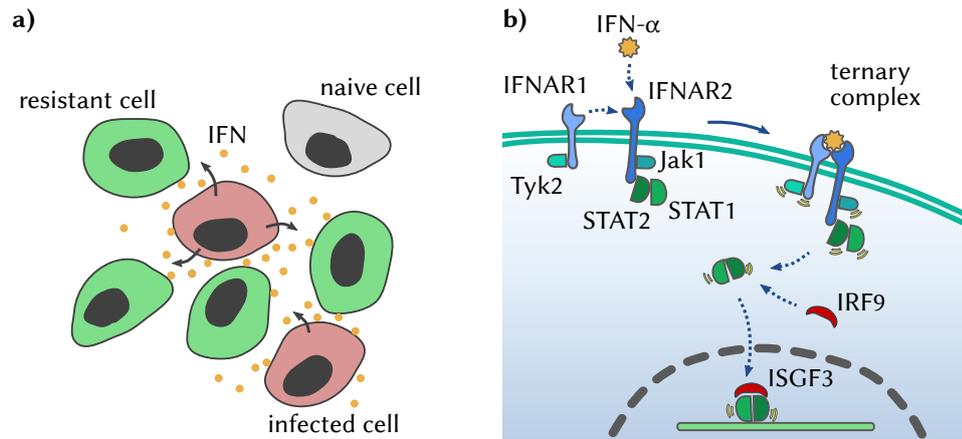


Figure 1.2.: Principles of the type I IFN response. **a)** Type I interferon is secreted by infected cells and induces an antiviral state in neighbouring cells, effectively inhibiting viral spread. **b)** The main signaling pathway of the cellular type I IFN response is the Janus kinase/signal transducers and activators of transcription (JAK/STAT) pathway. Type I interferons sequentially recruit interferon- α receptor chains 1 and 2 (IFNAR1 and IFNAR2) to form a ternary signaling complex. The receptor-associated kinases Janus kinase 1 (JAK1) and Tyrosine kinase 2 (Tyk2) phosphorylate each other, the receptor and the two downstream signaling molecules STAT1 and STAT2. These form activated heterodimers that complex with interferon regulatory factor 9 (IRF9) to form interferon-stimulated gene factor 3 (ISGF3), the main transcription factor of the type I IFN response. ISGF3 promotes the expression of IFN-stimulated genes (ISGs), which mediate an antiviral and antiproliferative state.

is proposed that IFN ligands first form a complex with IFNAR2 and then recruit IFNAR1 to form the active ternary signaling complex [42].

Neither of the two subchains of the IFN receptor has an intrinsic kinase function but rely on associated kinases for signal transduction. For IFNAR1 this is tyrosine kinase 2 (Tyk2) and for IFNAR2 this is Janus kinase 1 (JAK1). The formation of the ternary signaling complex allows the receptor-associated kinases to phosphorylate each other and the receptor, which creates binding and interaction sites for various downstream molecules (Fig 1.2 (b)). While multiple signaling pathways contribute to the IFN-dependent downstream signaling, the Janus kinase/signal transducers and activators of transcription (JAK/STAT) pathway is required for the antiviral and antiproliferative response (reviewed in [45]). Here, the downstream-signaling

is further mediated mostly by STAT1 and STAT2 that are supposed to preassemble at the IFNAR2 receptor chain [46, 47]. Upon activation by Jak1 and Tyk2, STAT1 and STAT2 form homo- and heterodimers that act as transcription factors. The major transcription factor in IFN signaling is interferon-stimulated gene factor 3 (ISGF3), which is composed of an activated STAT1:STAT2 dimer and interferon regulatory factor 9 (IRF9). ISGF3 is imported into the nucleus and binds to the IFN-stimulated response element (ISRE) within the promoter of IFN-stimulated genes (ISGs), promoting the expression of antiviral and antiproliferative genes [44]. The effector proteins coded in these genes restrict viral replication both by targeting viral processes as well as slowing cell growth. This includes inhibition of translation, induction of mRNA degradation or mediation of growth arrest [45].

1.3.2. Dynamics of IFN- α Receptor Dimerization

As mentioned, the initial event in the type I IFN signaling cascade is the formation of the ternary signaling complex. This formation has been well characterized and quantified on an ensemble level [42]. Besides the different affinities of IFNs towards the two IFN receptor chains, individual IFN subtypes also differ in their integral affinity towards the receptor, spanning about an order of magnitude. The strongest binding can be observed for IFN- β , which binds IFNAR1 with a 50 nM affinity and IFNAR2 with a 0.1 nM affinity, while IFN- $\alpha 2$ binds IFNAR1 with a 5 nM affinity and IFNAR2 with a 5 μ M affinity [48].

The different binding affinities of type I interferons towards both IFN-receptor chains do not only influence the assembly of the ternary signaling complex but also its stability. High affinity binding between the IFN receptor and its ligand stabilizes the ternary signaling complex and results in a longer lifetime. These different ternary complex lifetimes were proposed by Jacob Piehler and colleagues to be mainly responsible for the specific biological outcome that is evoked by different IFNs [42]. This hypothesis has been supported by Levin and co-workers who systematically altered the expression levels of both receptor chains and examined the biological effect of different IFNs under different receptor densities [49]. They could correlate the number of stabilized ternary signaling complexes with the quality of the biological outcome. While a small number of transiently formed ternary signaling complexes was sufficient to evoke an antiviral response, many long-term stable complexes were required for an antiproliferative response. Additionally, this decision seems to be binary for each cell. Considering that the expression level of the IFN- α receptor is very heterogeneous within most cell lines, there will only be a fraction of cells in each cell population that expresses the IFN receptor at levels high enough to evoke an antiproliferative response while most cells are capable of inducing an antiviral response [49].

The formation of the ternary signaling complex in live cells has been studied and described on a single-molecule level by Wilmes et al. [30]. This study could demonstrate the IFNAR1 and IFNAR2 do not predimerize but are only recruited into complex by an IFN ligand. When stimulated with IFN- α 2, about 70 % of IFNAR2/IFN- α 2 complexes succeed in recruiting IFNAR1 at a two-dimensional equilibrium dissociation constant K_D^T of 0.29 molecules/ μm^2 . These values were measured at room temperature using a sandwich of protein tags and quantum dots as fluorescent labels. In general, the group of Jacob Piehler has generated a large body of knowledge about single-molecule diffusion and interaction of the type I interferon signaling complex, that is very useful as a reference when establishing a new two-color single-molecule tracking system.

1.4. EGFR and c-Met Signaling in Non-Small Cell Lung Cancer

With around 1.5 million deaths every year, lung and bronchus cancer account for the highest cancer-related death toll worldwide [50]. A specific problem of lung cancer is its often late time of diagnosis and its early metastatic spread, which leave many patients inoperable already at the time of diagnosis. These patients rely on chemotherapy and additional targeted therapies like the treatment with tyrosine kinase inhibitors (TKIs), which inhibit signaling of the epidermal growth factor receptor (EGFR). Due to its extraordinary high mutation rate [51], lung tumors are fast to develop resistance against systemic treatment, reducing their benefit in survival to the range of months. Although many mediators of TKI resistance have been discovered, a full understanding of the cellular processes leading to TKI resistance is still missing. The most prominent drivers of resistance include mutations of EGFR and the amplification of the MET gene that causes overexpression of the hepatocyte growth factor receptor (HGFR or c-Met) [52]. Recently, a high ratio of c-Met vs. EGFR expression levels in non-small cell lung tumors was found to indicate *de novo* non-responders and correlate with a decreased progression free survival [53]. Hence, the interaction between EGFR and c-Met may be a driver of TKI resistance, that has not been well described so far.

1.4.1. The Role of EGFR and c-Met in Physiology and Pathophysiology

Both epidermal growth factor receptor (EGFR or Her1) and hepatocyte growth factor receptor (HGFR or c-Met) are growth hormone receptors and have important roles in cell survival, proliferation and differentiation. As such, EGFR and c-Met

signaling navigates embryonic development and tissue regeneration but also tumor progression.

EGFR is a member of the human epidermal growth factor receptor (Her) family that further includes Her2 (ErbB2), Her3 (ErbB3) and Her4 (ErbB4). Both receptors belong to the receptor tyrosine kinases. Hence, unlike the IFNAR receptors, EGFR and c-Met have an intrinsic kinase function. Upon ligand stimulation, EGFR and c-Met monomers typically form homodimers, which initiates the signaling cascade. Besides EGFR homodimers, also heterodimers between different Her receptors have been observed and discussed as a means of signal diversification [54]. EGFR binds seven different ligands that share a 35–40 amino acid EGF motif. Of these, EGF is one of the strongest binder and strongest inducer of homodimerization [55]. In humans, EGF is mostly produced in the kidney. In contrast, hepatocyte growth factor (HGF), which is expressed in many epithelial organs [56], is the only known ligand of c-Met. Ligand-induced dimerization activates the intrinsic kinase domain that in turn phosphorylates specific C-terminal tyrosine residues, which creates binding sites for downstream signaling molecules containing Src homology 2 (SH2) or phosphotyrosine binding (PTB) domains. While EGFR is generally assumed to exist as monomers in the absence of its ligands, it has also been proposed that EGFR can preform inactive EGFR homodimers, which are activated upon ligand binding [57].

The main signaling pathways of both EGFR and c-Met signaling are the mitogen activated protein kinase (MAPK) pathway and the 3-phosphoinositide-dependent protein kinase/protein kinase B (PI3K/Akt) pathway that are activated through the recruitment of the adapter proteins Grb2 and Gab1 [58]. Along the MAPK pathway, the cellular signal is transmitted by the small GTPase Ras and amplified by sequential activation of the kinases Raf1, MEK and ERK(1/2). ERK interacts with a variety of target proteins that drive cell proliferation, differentiation and development [59]. The PI3K/Akt pathway is initiated by the activation of PI3K, a kinase that catalyzes the formation of the second messenger molecule phosphatidylinositol-3,4,5-trisphosphate (PIP3). PIP3 in turn activates protein kinase B (PKB or AKT), which again interacts with a large number of downstream target proteins. In general, these effector proteins promote cell survival and motility [56]. When overactive, both pathways facilitate the growth and spread of tumors. Consequently both EGFR signaling as well as c-Met signaling is often altered in lung cancer.

1.4.2. EGFR as a Target in Non-Small Cell Lung Cancer Therapy

In the developed world, non-small cell lung cancer (NSCLC) is the predominant form of lung cancer [60]. Mutations of the epidermal growth factor receptor (EGFR) belong to the most frequent driver mutations found in NSCLC patients [61]. In

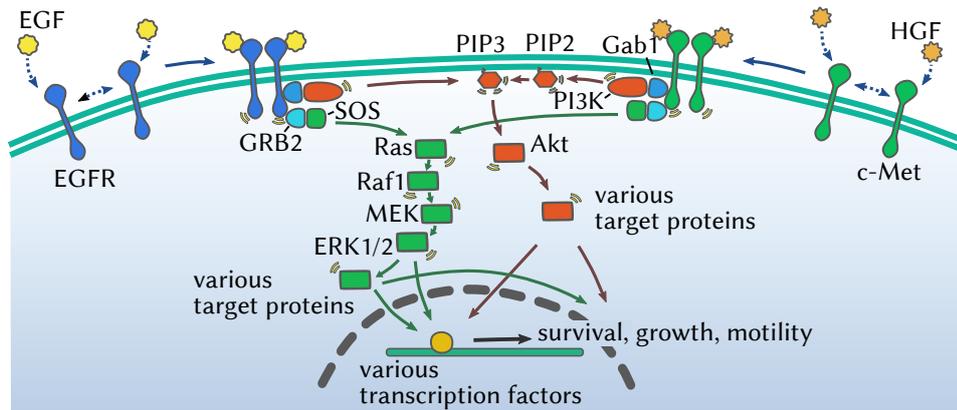


Figure 1.3.: General pathways of EGFR and c-Met signaling. Epidermal growth factor receptor (EGFR) and hepatocyte growth factor (c-Met) belong to the family of receptor tyrosine kinases and form homodimers upon stimulation with their respective ligands EGF and HGF. This dimerization activated the kinase domain of the receptors that phosphorylate each other on the receptor, creating binding sites for the adapter molecules GRB2 and Gab1. These adapter molecules mediate the activation of the mitogen-activated protein kinases (MAPK) pathway and the phosphatidylinositol 3-kinase/protein kinase B (PI3K/Akt) pathway. Both pathways promote survival, cell growth and motility.

10–30 % of patients, a L858R point mutation or an inframe Δ E746-A750 deletion leads to an overactivation of EGFR [61], which promotes cell proliferation and cancer progression. This effect is further enhanced by an overexpression of EGFR that can be found in many patients [62].

To delay cancer progression, tyrosine kinase inhibitors (TKIs) are used as first, second or subsequent line therapy in advanced, unresectable NSCLC [63]. The first generation TKIs Erlotinib (Tarceva[®]) and Gefitinib (Iressa[®]) represent ATP-analogues that competitively bind the ATP binding pocket of EGFR in order to inhibit EGFR signaling. In contrast to this reversible binding mode, second and third generation TKIs Afatinib (Gilotrif[®]), WZ4002, AZD9291 or CO-1686 covalently bind to Cys797 in the active site of EGFR [64, 65]. For patients with activating EGFR mutations, TKI treatment significantly delays disease progression and shows a better response rate and quality of life compared to first-line chemotherapy [63]. Monoclonal antibodies targeting EGFR are approved in combination with chemotherapy (Cetuximab) or still under investigation (Panitumumab, Nectinmumab or MM-151)

Unfortunately due to the high mutation rate of lung tumors, almost all patients acquire resistance against TKIs over time [63]. Resistance to first generation TKIs typically appears after 9 to 14 months [65] and in more than half of TKI resistant patients, a secondary T790M mutation in EGFR can be found [66]. By recovering the ATP binding affinity of EGFR, this mutation reduces the efficiency of the competitive TKIs [67]. The inhibitory effects of second or third generation TKIs are eventually overcome by a C797S mutation in EGFR [68]. This removes the covalent binding partner of the inhibitors and thus turns the mutated EGFR resistant.

An additional mechanism that promotes TKI resistance is the amplification of the MET gene that is observed in 5–20 % of TKI resistant patients [66,69]. It is proposed that amplification of c-Met signaling can bypass and compensate for reduced EGFR signaling in promoting tumor growth due to the overlap of c-Met and EGFR signaling pathways. However, the exact mechanism of c-Met-mediated resistance is unknown. Besides their interconnected signaling pathways, recent studies indicate that a direct interaction between EGFR and c-Met might play a role in TKI sensitivity and resistance. Direct interaction of EGFR and c-Met have been observed in different cell lines positive for MET amplification [70]. In a relatively recent study, TKIs were reported to decrease the amount of EGFR homodimers but not EGFR/Her2 heterodimers in three esophageal cells [71]. Although not included in that study, it can be speculated that also previously poorly described heterodimerization pairs like EGFR/c-Met may profit from a similar specific stabilization. This assumption is supported by the recent description of a bispecific EGFR/c-Met antibody that induced a complete and durable regression of human lung xenograft tumors when combined with a third generation TKI [72]. In another study, the progression free survival of NSCLC patients was found to correlate with the expression ratio of c-Met and EGFR [53]. Here, a high c-Met/EGFR ratio indicated *de novo* TKI non-responders and a reduced progression free survival. Conversely, even when c-Met expression is high, a similarly high expression of EGFR may restore TKI sensitivity. Taken together, these studies indicate that receptor-receptor interactions may be a crucial regulation point and that EGFR/c-Met heterodimerization in particular may play an important role in altering TKI sensitivity.

In order to understand the role of EGFR and c-Met within the signaling of NSCLC cells and the emergence of TKI resistance, the labs of Ursula Klingmüller¹ and Jens Timmer² have developed a dynamic pathway model on combined EGFR and c-Met signaling based on time resolved quantitative western blot data. This

¹German Cancer Research Center, Heidelberg

²Freiburg University

model predicted a crucial role for EGFR/c-Met heterodimers in the modulation of cell signaling and sensitivity toward TKIs.

1.5. Scope of this Study

This study aims to establish a two-color, single-molecule tracking system in order to detect and quantify receptor-receptor interactions in two clinically relevant biological systems. Being already well described, the formation of the IFN- α signaling complex will serve as a reference system. In contrast to most published data, I will study the formation of the IFN- α signaling complex at physiological more relevant conditions using relatively light labels and performing experiments at 37 °C. In addition, the diffusion and interaction dynamics of the type I interferon receptor will be combined in a comprehensive stochastic reactive diffusion model in cooperation with Nikolas Schnellbacher of the group of Ulrich Schwarz³. The formation of heterodimers between the two growth factor receptors EGFR and c-Met has been hypothesized by the group of Ursula Klingmüller to influence the sensitivity of lung cancer cells towards tyrosine kinase inhibitors. Based on the developed two-color tracking system, this study will directly test the EGFR/c-Met heterodimerization hypothesis and describe the interaction of EGFR and c-Met upon stimulation with their respective ligands.

With a very recent commercial Nikon Ti Eclipse TIRF microscope at hand, this study will profit from a robust, versatile and single-molecule sensitive imaging system. As a first step, a labeling strategy based on SNAPf-tag and HaloTag fusion constructs will be established and characterized. As a crucial information for quantitative microscopy, the degree of labeling (DOL) of both protein tags will be determined on a single-molecule level for varying live-cell labeling conditions. For this, a modular calibration construct based on a fusion between both protein tags and an additional enhanced green fluorescent protein (eGFP) will be designed, cloned and stably expressed in different cells lines. In order to determine the DOL of each tag based on its colocalization with the eGFP signal, an automated colocalization analysis, including particle detection, image registration and colocalization detection, will be developed and evaluated based on simulated single-molecule data. Besides estimating the DOL of SNAPf-tag and HaloTag for a given labeling condition, the aim of this part of the study is also to generally evaluate the suitability of both tags for TIRF-based single-molecule microscopy and to find their optimal labeling conditions and their limitations.

Next, a two-color-tracking routine will be established and used to detect and quantify receptor-receptor interactions in live cells. This part of the study will be

³Heidelberg University

accomplished in collaboration with Ursula Klingmüller⁴, Karl Rohr³ and Ulrich Schwarz³. To this end, the receptors of both biological systems will be genetically fused to either SNAPf-tag or HaloTag and the respective receptor pairs will be simultaneously tracked under different stimulation conditions. In respect to the IFN- α receptor complex, a general and systemic understanding of the parameters that govern ternary complex formation will be promoted by combining two-color tracking with an stochastic single-molecule model of the diffusion and interaction of the two receptor subchains. Such an in-depth understanding will help to identify facilitators and inhibitors of ternary complex formation, which may influence the efficiency of IFN- α treatment. In terms of the EGFR/c-Met system, the hypothesis of EGFR/c-Met heterodimerization will be tested. In addition, the frequency and stability of these heterodimers will be determined in dependence on different ligands. This will help to understand how the interaction between EGFR and c-Met may influence the sensitivity of lung cancer cells toward tyrosine kinase inhibitors.

⁴German Cancer Research Center, Heidelberg

Part II.

Experiments and Results

Chapter 2.

The Degree of Labeling of SNAPf-tag and HaloTag

In order to find the best labeling conditions for SNAPf-tag and HaloTag, it is necessary to determine their respective degree of labeling (DOL), that is the fraction of tags that is actually labeled with a fluorescent dye. Furthermore, the DOL is a crucial correction factor for the quantification of molecule numbers and protein concentrations from microscopy data. As mentioned in chapter 1.2.3 on page 7 of the introduction, the DOL of a protein tag can be estimated on a single-molecule level by a colocalization analysis with an additional, independent signal that is however associated with the tag. Wilmes et al., for example, used a direct fusion between SNAP-tag and HaloTag and cross-referenced both signals against each other in order to estimate the DOL of each tag [30]. Latty et al. introduced an additional antibody signal against which the protein tags could be referenced [31]. Both methods however, suffer from unspecific labeling of the reference signal (e.g. antibodies or tag substrates that bind unspecifically to the cell membrane or the glass surface). As these unspecific signals do not preferably colocalize with the protein tag signal, the DOL may easily be underestimated. To solve this problem, I have developed an DOL calibration probe that uses an enhanced green fluorescent protein (eGFP) as a practically background free reference signal. Due to its modular design, this probe can be easily equipped with different protein tags and expressed in various cell lines.

To determine the DOL of SNAPf-tag and HaloTag, the DOL calibration probe consisted basically of an enhanced green fluorescent protein (eGFP) fused to both protein tags and equipped with a membrane anchor to allow for TIRF microscopy in cells (see figure 2.1 (a)). In addition, I introduced a 5.7 nm long α -helical linker domain [73] to avoid energy transfer between the protein tags and a C-terminal His-tag for further antibody staining or purification, if needed. Using spectrally separate dye substrates for SNAPf-tag and HaloTag, this GFP-extended Staining Efficiency Probe (gSEP) allows to determine the DOL of both tags in one combined three-color experiment.

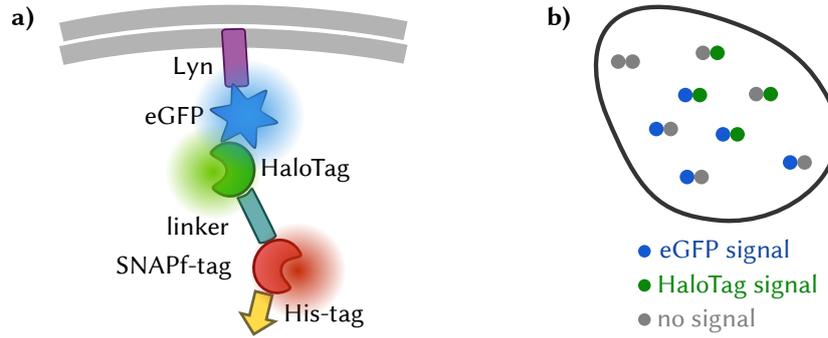


Figure 2.1.: Principle of the GFP-enhanced Staining Efficiency Probe (gSEP). **a)** The construct is composed of a fusion between eGFP, HaloTag and SNAPf-tag that is further equipped with a N-terminal membrane anchor (Lyn), an α -helical linker between the two protein tags and a C-terminal His-Tag. **b)** The DOL can be determined as the fraction of double-labeled constructs (blue and green for eGFP and HaloTag, respectively) from all eGFP-positive constructs. Based on this example, the DOL would be $\frac{2}{5}$.

The determination of the DOL is based on a colocalization analysis in fixed cells on a single-molecule level. Since eGFP signal is basically background-free it can be used to reliably localize the gSEP constructs within the cell membrane. Assuming that the signal from eGFP and from either protein tag are independent, the DOL can be calculated as the number of double labeled constructs, i.e. constructs that show both an eGFP signal as well as a signal from the respective tag, divided by the number of constructs that show at least an eGFP signal (see equation 2.1 and Figure 2.1 (b)).

$$DOL_{tag} = \frac{|tag \cap eGFP|}{|eGFP|} \quad (2.1)$$

,where $eGFP$ is the set of particles in the eGFP channel and tag is the set of particles in the channel of the respective protein tag. Applying this formula to the situation depicted in Figure 2.1 (b) would yield a DOL of $\frac{2}{5} = 40\%$.

In the following chapter I want to show how the GFP-extended Staining Efficiency Probe was cloned, stably expressed in cells and finally used to determine the DOL of SNAPf-tag and HaloTag for 40 different staining conditions.

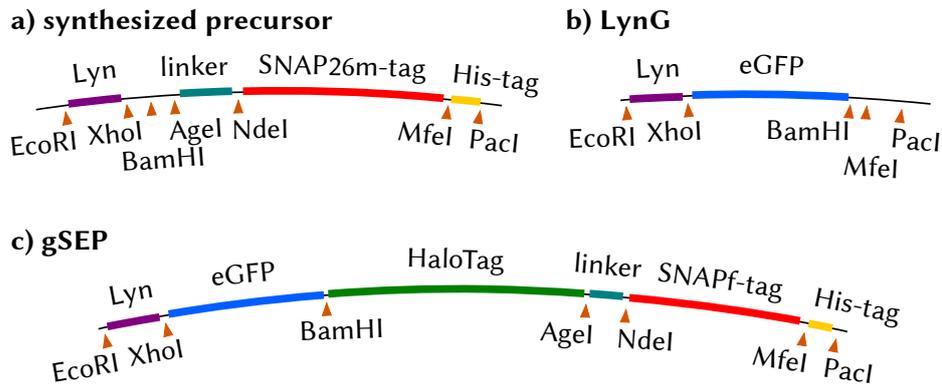


Figure 2.2.: Architecture of the GFP-extended Staining Efficiency Probe (gSEP) and its negative control (LynG). gSEP was stepwise cloned from a synthesized precursor that contained all short domains and a set of unique cloning sites (a). To create gSEP, the larger domains eGFP and HaloTag were added to the precursor and SNAP26m was replaced by SNAPf (c). A truncated form of gSEP, that exists of eGFP and the N-terminal myristoylation domain of Lyn kinase, represents a negative staining control (b).

2.1. Cloning of the GFP-extended Staining Efficiency Probe (gSEP)

I ordered a modular precursor of the Staining Efficiency Probe as a synthetic gene from integrated DNA Technologies, which consisted of a fusion between the N-myristoylation/palmitoylation domain of Lyn kinase [74] (membrane anchor), a 5.7 nm long α -helical linker domain [73], SNAP26m-tag [20] and a six amino acid His-tag (see Figure 2.2 (a)). The precursor further included seven unique restriction sites to allow for the insertion of two additional domains between the membrane anchor and the α -helical linker as well as the easy removal or exchange of each domain. As this design contained all of the smaller domains of the final construct, the necessary cloning effort was largely reduced compared to *de novo* cloning. At the same time, manually adding the remaining larger domains lowered the total size of the synthesized precursor to 730 bp. Since the synthesis of synthetic genes was charged by the number of base pairs, the ordered construct was a compromise between costs and remaining cloning effort.

The gSEP construct was stepwise assembled from the ordered precursor in the following six cloning steps. After each step, the respective product was verified by

sequencing. The cloning process is shown in more detail in section A.1 on page 129 of the appendix.

The gSEP precursor was delivered in a pIDTSmart vector with ampicillin resistance and was first transferred into an empty retroviral pMOWS vector using EcoRI and PacI restriction sites. Next, I used BamHI and AgeI restriction sites to insert HaloTag. For this, I amplified the HaloTag sequence by PCR from a pMOWS-Tyk2-Halo plasmid provided by Florian Salopiata from the group of Ursula Klingmüller¹ as a template as well as PCR primers with a respective BamHI and AgeI overhangs^{2,3}.

In a next step, I replaced SNAP26m-tag with SNAPf-tag, which is reported to have a faster reaction kinetic towards its dye substrate *in vitro* [23] and *in vivo* [75]. SNAPf-tag was bought from New England Biosciences and amplified with NdeI and MfeI restriction site overhangs via PCR^{4,5}. SNAP26m-tag was then replaced by SNAPf-tag using NdeI and MfeI restriction sites.

At last, I inserted eGFP between the XhoI and BamHI restriction sites. For this, eGFP was amplified via PCR with XhoI and BamHI overhangs using a HaloTag-eGFP-Tubulin plasmid that was provided by Klaus Yserentant from the groups of Dirk-Peter Herten⁶ and Rasmus Schröder⁶ as a template^{7,8}.

One decisive parameter for the single-molecule microscopy-based determination of the DOL of SNAPf-tag and HaloTag is the molecular density of the gSEP construct within the plasma membrane. While high molecule densities eventually provide better statistics, the density must be sufficiently low to detect single, isolated molecules in a diffraction-limited image. As the retroviral pMOWS plasmid induces a high expression [76], I decided to transfer the gSEP construct into the lower expressing retroviral pBABE plasmid [77]. Thanks to similar multiple cloning sites of pMOWS and pBABE, the full gSEP sequence could be extracted from pMOWS-gSEP via digestion with EcoRI and PacI and inserted in the pBABE plasmid using the same restriction sites.

While eGFP provides a background-free label, staining the protein tags can introduce unspecific background signal [24], for example when dye substrate binds to off-target structures like lipids or the glass surface. In order to estimate the amount of unspecific labeling for different staining conditions, I created a truncated

¹German Cancer research Center, Heidelberg

²fwd: 5'-CGCGGATCCGCAGAAATCGGTACTGGCTTCCATTC-3'

³rev: 5'-CGCACCGGTGCCGAAATCTCCAGCGTCGAC-3'

⁴fwd: 5'-GCGCATATGGACAAAGACTGCGAAATGAAGCGCACC-3'

⁵rev: 5'-GCGCAATTGACCCAGCCCAGGCTTGCCCAG-3'

⁶Heidelberg University

⁷fwd: 5'-ATCCCTCGAGACCATGGTGAGCAAG-3'

⁸rev: 5'-CACGGATCCCTGTACAGCTCGTC-3'

version of gSEP by removing both protein tags via digestion with BamHI and MfeI, blunting the sticky ends and ligating the plasmid again. The resulting construct consisted of the N-terminal myristoylation domain of Lyn kinase followed by eGFP (LynG, see Figure 2.2 (b)).

2.2. Creation of Stable Cell Lines

In order to establish stable cell lines, the validated pBABE-gSEP and pBABE-LynG plasmids were each virally transduced into Huh7.5 cells using the Phoenix Ampho system as described in section 6.2.1 on page 113 and stable cells were selected with puromycin at a concentration of $1.5 \mu\text{g mL}^{-1}$. The resulting cell lines will be referred to as Huh7.5-pBABE-gSEP and Huh7.5-pBABE-LynG. To test the expression and functionality of both constructs, HaloTag and SNAPf-tag were simultaneously stained with HaloTag tetramethylrhodamine ligand (HaloTag TMR ligand) and siliconrhodamine-benzylguanine (SiR-BG), respectively and signals in the corresponding spectral channels as well as in the eGFP channel were recorded using single-molecule sensitive total internal reflection fluorescence (TIRF) microscopy. In short, cells of each cell lines were seeded in a cleaned LabTek, stained with 1 nM HaloTag TMR ligand and 1 nM SiR-BG overnight at 37°C and 5% CO_2 and subsequently washed and fixed as described in section 6.7.3 on page 119. The signal of the three spectral channels (blue, green and red for emission of eGFP, TMR and SiR, respectively) was sequentially recorded from low to high energies, i.e. from red to blue, in order to avoid prebleaching. Figure 2.3 shows the signal of each spectral channel from an exemplary cell of each cell line.

The presence of single-molecule signals under TIRF conditions indicated the functionality of the membrane anchor. If the membrane anchor had been non-functional, the constructs would have been expressed to the cellular cytosol and washed out during fixation. As expected, eGFP signal was present in Huh7.5-pBABE-gSEP cells as well as in Huh7.5-pBABE-LynG at a density that is well suitable for colocalization analysis on a single-molecule level. By trend, eGFP signal was brighter in Huh7.5-pBABE-LynG cells. This could indicate Förster resonance energy transfer (FRET) between eGFP and TMR, which is further examined in the following chapter 2.3. A specific TMR and SiR signal could only be detected in Huh7.5-pBABE-gSEP cells, which indicates that HaloTag and SNAPf-tag were efficiently and specifically labeled. In contrast, Huh7.5-pBABE-LynG cells showed a limited amount of autofluorescence as well as unspecific staining that appeared stronger in the SiR channel than in the TMR channel.

These results indicate that all components of the gSEP and the LynG construct were expressed and functional. In addition, the observed expression level was

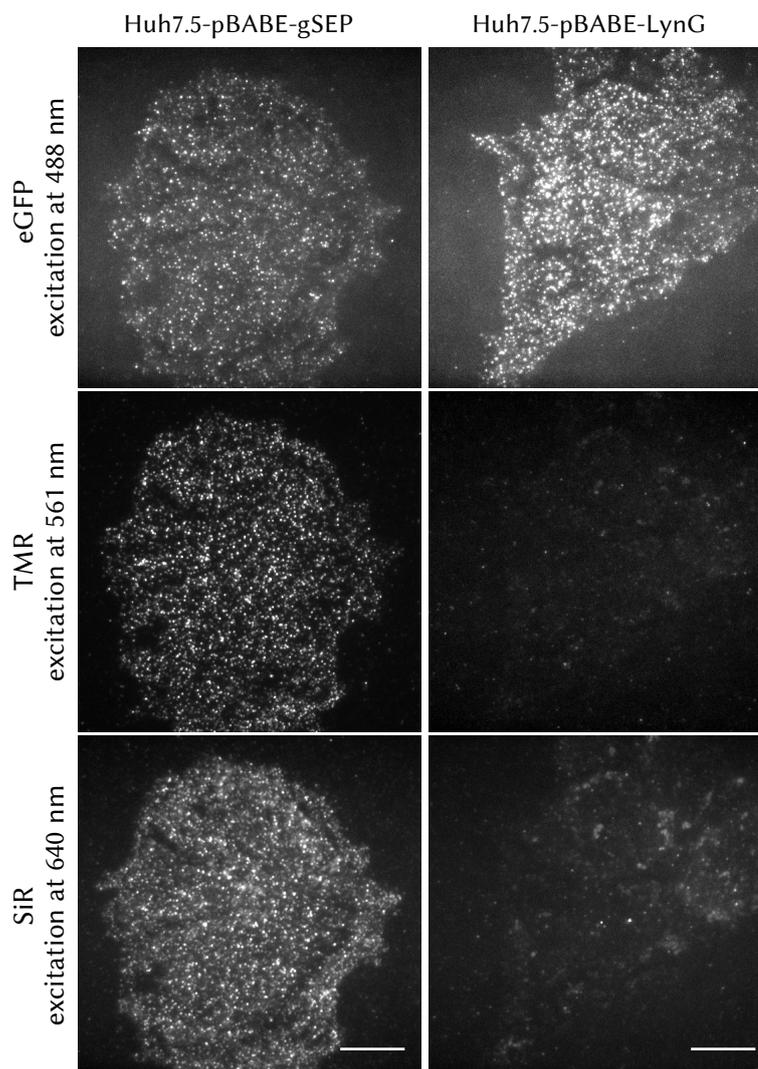


Figure 2.3.: Expression of gSEP and LynG constructs in stable Huh7.5 cell lines. While eGFP signal could be detected in both cell lines, only Huh7.5-pBABE-gSEP cells showed a specific signal in the TMR channel and the SiR channel upon overnight live cell labeling with 1 nM HaloTag TMR ligand and 1 nM SiR-BG, indicating specific labeling of HaloTag and SNAPf-tag. Images represent the temporal average of frame 10 to 20 of a 100 frame movie. For better comparison between the cell lines, all images of the same channel have the same intensity scale. Scale bar is 10 μ m.

assumed low enough for single-molecule detection and high enough to produce reasonable statistics. For a quantification of eGFP particle densities in Huh7.5-pBABE-gSEP and Huh7.5-pBABE-LynG cells see Figure 2.5 on page 34.

2.3. Determination of Förster Resonance Energy Transfer between the Fluorophore Pairs in the gSEP Construct

Förster resonance energy transfer (FRET) is a radiation free energy transfer between a donor fluorophore and an acceptor fluorophore (see also section 1.2.1 on page 4 of the introduction). Instead of emitting a fluorescence photon, an excited donor fluorophore can transfer its energy to an acceptor fluorophore, which is put in an excited state and can emit a fluorescent photon of its specific wavelength. In respect to the colocalization analysis for the determination of the DOL, a high FRET efficiency would violate the requirement that all the observed signals are independent of each other. Due to the energy transfer, donor fluorophores emit fewer photons, which makes them dimmer and harder to detect. At high efficiencies, FRET would therefore reduce the number of observed particles in all channels where the respective fluorophores could serve as donor fluorophores when these donor fluorophores are colocalized with their respective acceptor fluorophore. As a consequence, colocalized particles will to some extent be misclassified as acceptor-only particles. According to equation 2.1 on page 24, reducing the number of colocalized particles will directly influence the calculated DOL. Therefore, a valid DOL can only be determined if FRET efficiencies between the fluorophore pairs are kept small.

Within the gSEP construct, FRET can in principle occur between eGFP and the HaloTag substrate as well as between the HaloTag substrate and the SNAPf-tag substrate. The efficiency of FRET between two fluorophores depends strongly on their distance and can effectively be reduced by spatially separating donor and acceptor fluorophore [73]. While such a separation is achieved between HaloTag substrate and SNAPf-tag by the introduced α -helical linker, the direct fusion of GFP to HaloTag could promote FRET between these two domains.

Besides the distance between two fluorophores, also their spectral overlap critically influences FRET efficiency. For all experiments in this dissertation, HaloTag was labeled with HaloTag tetramethylrhodamine ligand (HaloTag TMR ligand) and SNAPf-tag was labeled with siliconrhodamine-benzylguanine (SiR-BG). Hence the possible FRET pairs for the gSEP constructs are eGFP (donor)-TMR (acceptor) and TMR (donor)-SiR (acceptor). In principle, reversing the dye substrates, i.e. labeling HaloTag with SiR and SNAPf-tag with TMR would effectively disrupt FRET be-

tween eGFP and TMR due to the increased distance. However, the estimated DOLs were initially planned to support the two-color tracking experiments described in chapter 3 and chapter 4. In these experiments, SNAPf-tag and HaloTag were labeled with SiR-BG and HaloTag TMR ligand, respectively, as these combinations proved most efficient. To guarantee consistency, I decided to maintain these labeling pairs also for the determination of the DOLs.

To experimentally determine the FRET efficiency between eGFP and HaloTag-bound TMR (Halo-TMR) as well as between HaloTag-bound TMR and SNAPf-tag-bound SiR (SNAP-SiR), single-molecule alternating laser excitation (smALEX) measurements of both fluorophore pairs were performed as described in chapter 6.5 on page 116 of material and methods. In short, live Huh7.5-pBABE-gSEP cells were stained overnight with either 1 nM HaloTag TMR ligand alone or with 1 nM HaloTag TMR ligand and 1 nM SiR-BG. All cells were washed three times with full growth medium, fixed with 4 % paraformaldehyde and imaged on a TIRF microscope at room temperature. Cells that were stained with HaloTag TMR ligand only were used to determine the FRET efficiency between eGFP as donor and Halo-TMR as acceptor. Cells that were stained with HaloTag TMR ligand and SiR-BG were used to determine the FRET efficiency between Halo-TMR as donor and SNAP-SiR acceptor.

A total of 2308 colocalized eGFP and TMR signals from 13 individual cells were used to quantify the FRET efficiency between eGFP and Halo-TMR (Fig 2.4, left column), while the FRET efficiency between Halo-TMR and SNAP-SiR was quantified based on 1445 colocalized TMR and SiR signals from 8 individual cells (Fig 2.4, right column). The FRET efficiency between eGFP and Halo-TMR averaged 0.47, which represented a moderate FRET efficiency. However, with a standard deviation of 0.35, the underlying distribution was very broad. This could be an effect of the free movement and rotation of the two domains towards each other that influences the distance and orientation of the fluorescent dipoles and therewith the FRET efficiency. With an average value of 0.52, the determined stoichiometry between eGFP and Halo-TMR confirmed the expected 1:1 ratio between the two domains and indicated that the chosen measuring time of 5 s was compatible with the average bleaching time of the fluorophores. The E-S histogram of this FRET pair showed a slight overrepresentation of the donor-only population at a stoichiometry of $S = 1$ and a FRET efficiency of $E = 0$. In fact, this is an indicator for efficient FRET. At high FRET efficiencies, acceptor fluorophores can be excited directly by the excitation laser and indirectly by the donor fluorophore. This additional excitation path leads to a faster bleaching of the acceptor fluorophore and an overrepresentation of the remaining donor-only population over time.

Between Halo-TMR and SNAP-SiR, a considerably lower FRET efficiency of

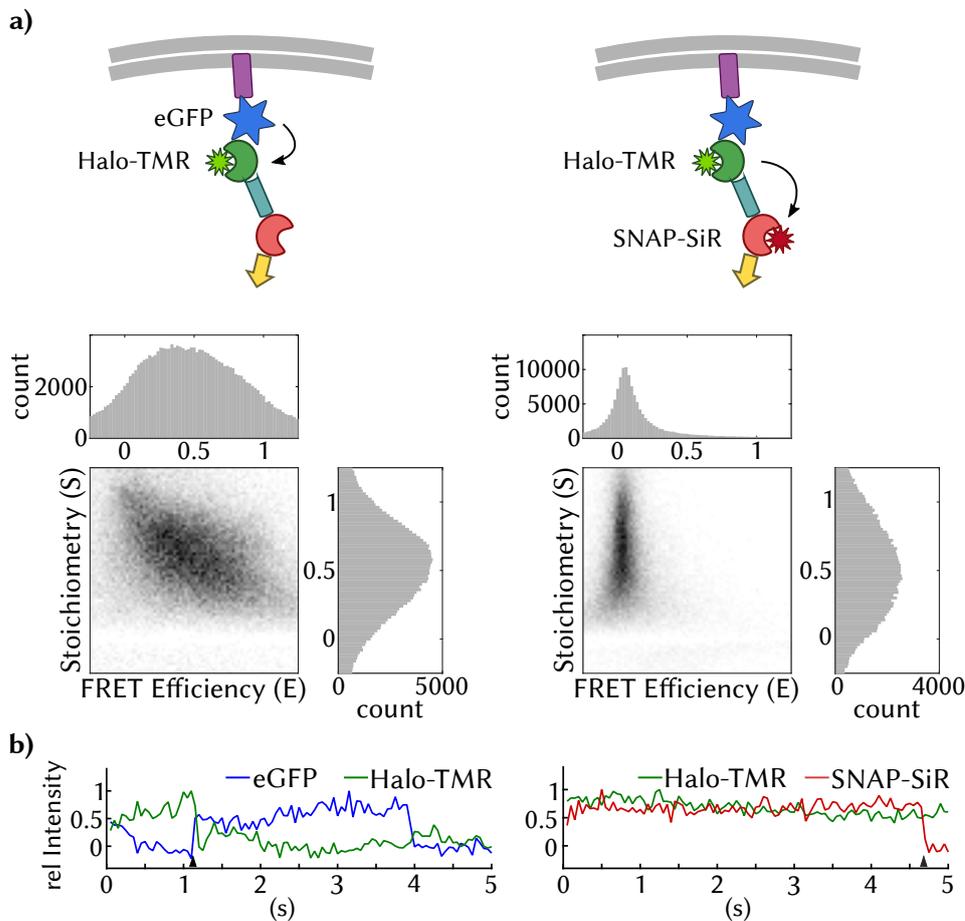


Figure 2.4.: Determination of the FRET efficiency and stoichiometry between the two fluorophore pairs. a) Single-molecule alternating laser excitation (smALEX) measurements revealed a widely distributed FRET efficiency between eGFP and HaloTag-bound TMR (Halo-TMR) with a mean efficiency of 0.47. In contrast, Halo-TMR and SNAPf-tag-bound SiR (SNAP-SiR) showed a very low FRET efficiency with a mean of 0.12. With a stoichiometry around 0.5, both fluorophore pairs confirmed the 1:1 ratio between the domains. **b)** Single-molecule intensity trajectories of a colocalized pair of eGFP and Halo-TMR over time showed an increased eGFP (donor) intensity when Halo-TMR (acceptor) was bleached (left, acceptor bleaching event marked by a dark triangle). For a colocalized pair of Halo-TMR and SNAP-SiR, no change in Halo-TMR (donor) intensity was observed upon SNAP-SiR (acceptor) bleaching (right). The traces show the relative and background-corrected intensities of each signal.

0.12 could be measured. With a standard deviation of 0.22, the distribution was also considerably sharper than the FRET efficiency distribution of eGFP and Halo-TMR. This confirmed the FRET-reducing effect of the α -helical linker. The mean stoichiometry of 0.51 again confirmed the expected 1:1 ratio between the two protein tags.

The difference in FRET efficiencies of the two fluorophore pairs could also be observed on a single-molecule level. Background-corrected single-molecule intensity traces of colocalized eGFP and Halo-TMR signals revealed an increase in donor intensity upon acceptor bleaching, while such a behavior could not be detected for colocalized Halo-TMR and SNAP-SiR pairs (Fig 2.4 (b), acceptor bleaching is marked with a black triangle).

As a summary, using smALEX measurements I could reveal considerable FRET between eGFP and HaloTag-bound TMR within the stained gSEP construct that might influence the determination of the DOL of HaloTag and SNAPf-tag. In contrast, FRET between Halo-TMR and SNAP-SiR is largely prevented by the introduced α -helical linker.

2.3.1. Acceptor Bleaching to Avoid FRET between eGFP and TMR Within the gSEP Construct

As a result of the above described smALEX measurement, FRET between eGFP and HaloTag-bound TMR (Halo-TMR) within the gSEP construct was identified as a possible obstacle for the determination of the DOL of the protein tags. Since recloning the gSEP construct was considered impractical, I chose acceptor photobleaching as an alternative way to circumvent FRET between eGFP and Halo-TMR. For photobleaching, high laser powers are used to irreversibly destroy the acceptor fluorophores, making them incapable of accepting energy via FRET. Practically, the acceptor channel is imaged and afterwards photobleached. Now the donor channel can be imaged without any occurrence of FRET as no functional FRET acceptors are present any more. For the determination of the DOL however, this approach is only justifiable if photobleaching of the acceptor fluorophores (Halo-TMR) does not considerably prebleach the donor fluorophores (eGFP) at the same time as this would artificially reduce the eGFP population. With regard to equation 2.1, this would at first sight only reduce the size of the statistical sample without necessarily changing the ratio between the number of particles in the eGFP channel and the number of colocalized particles. However under FRET conditions, eGFP may likely bleach slower when colocalized with TMR because FRET represents an additional nondestructive pathway to depopulate the excited state. Hence, the ratio between the number of particles in the eGFP channel and the number of colocalized particles would be shifted towards colocalized particles, which would

lead to an overestimation of the DOL. Under this consideration, it is important to estimate the degree of eGFP prebleaching induced by TMR bleaching in order to evaluate if acceptor photobleaching is a suitable method to avoid FRET between eGFP and Halo-TMR within the gSEP construct.

To estimate the required exposure time to thoroughly photobleach Halo-TMR as well as the influence of this photobleaching step on the number of eGFP particles, I recorded photobleaching curves of Halo-TMR and eGFP in fixed cells. For this, I seeded Huh7.5-pBABE-gSEP cells in a cleaned Labtek and stained them overnight with 10 nM HaloTag TMR ligand or left them unstained. All cells were washed three times with full growth medium, fixed with 4 % paraformaldehyde and imaged on a TIRF microscope at room temperature. Photobleaching of TMR was performed with a 561 nm laser at a power of 8.1 mW (power of the laser beam as it leaves the objective). For each frame of the recorded movies, single particles were detected using the Point Source Detection algorithm provided by the Matlab-based tracking software u-track [36] as described in section 2.4 on page 35.

At such a high laserpower, Halo-TMR showed a rapid photobleaching with half of the number of particles being lost after around 5 s already (Fig 2.5 (a)). A thorough photobleaching of Halo-TMR could be reached after around 2 min, when only 2.5 % of the initial number of particles were left. Hence, a photobleaching duration of 2 min was chosen for all further Halo-TMR photobleaching steps.

To determine the effect of TMR bleaching on eGFP particles, I observed the number of eGFP particles over time in unstained Huh7.5-pBABE-gSEP cells, i.e. where eGFP was the only fluorophore present, at a laserpower of 1.8 mW (power of the laser beam as it leaves the objective) using a 488 nm laser. To estimate the effect of TMR photobleaching on eGFP, I stopped eGFP observation after 2.5 s, performed a 2 min TMR photobleaching step and then continued eGFP observation. The photobleaching kinetic of eGFP was unchanged when the 2 min acceptor bleaching step was introduced (compare Fig 2.5 (b) and Fig 2.5 (c)). Most importantly, the number of detected eGFP particles before and after the TMR bleaching step was largely unchanged, i.e. eGFP molecules are not photobleached when exposed to high laserpowers at 561 nm. In contrast, a 1 min photobleaching step with 4.7 mW at 488 nm effectively photobleached eGFP (Fig 2.5 (d)). In Huh7.5-pBABE-gSEP cells where both eGFP and Halo-TMR were present, photobleaching of Halo-TMR nearly doubled the number of detected eGFP particles (Fig 2.5 (e)), which can be explained by acceptor bleaching under FRET conditions. This masking effect of eGFP particles by Halo-TMR particles was an additional indication of efficient FRET between eGFP and Halo-TMR within the gSEP construct and further supported the introduction of a TMR bleaching step.

To compare the expression levels of the gSEP construct and the LynG construct,

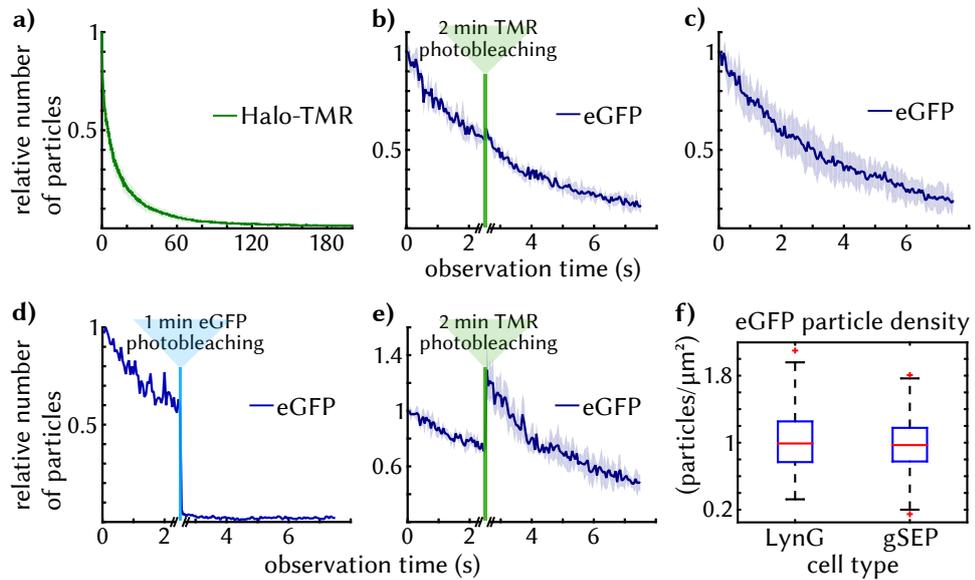


Figure 2.5.: Halo-TMR bleaching and its influence on eGFP intensity. a)

With a laserpower of 8.1 mW at 561 nm, the number of detected HaloTag-bound TMR (Halo-TMR) particles was halved after 5 s while only 2.5 % of the initial number of detected particles was left after 2 min. **b)** A 2 min Halo-TMR photobleaching step (indicated as a green vertical line) did not notably change the number of detected eGFP particles. The decrease of detected eGFP particles over time was similar to a continuous observation of eGFP as shown in **c)**, where no photobleaching was performed. **d)** A 1 min photobleaching step with 4.7 mW at 488 nm (indicated as a blue vertical line) resulted in a complete photobleaching of eGFP. **e)** When both eGFP and Halo-TMR were present within the gSEP construct, photobleaching of Halo-TMR strongly increased the number of detected eGFP particles due to acceptor bleaching. With exception of **d)**, all graphs show the mean number of detected particles from three different cells and its standard deviation. The data in **d)** is based on a single measurement. **f)** While the density of eGFP particles showed a large variation between individual cells, its mean value and distribution were well comparable for the two cell types, which expressed either the gSEP construct or the LynG construct. The density of eGFP particles for both cell types was determined after a 2 min Halo-TMR photobleaching step for 421 Huh7.5-pBABE-gSEP cells and 406 Huh7.5-pBABE-LynG cells.

the density of eGFP particles was determined for 421 Huh7.5-pBABE-gSEP cell images and 406 Huh7.5-pBABE-LynG cell images (Fig 2.5 (f)). For this, the cell area in each image was obtained via segmentation of the eGFP channel using a custom-written algorithm in Fiji [78]. The segmentation process included a rolling ball algorithm with a radius of 20 pixel to extract local average intensities followed by a low-pass filter to eliminate features smaller than 25 pixel, a median filter with a radius of 30 pixel to homogenize the intensity throughout the cell and a final thresholding based on Fiji's Triangle algorithm. Within this segmented area, eGFP particles were detected and the number of detected eGFP particles was divided by the corresponding cell area. Although no HaloTag labeling was performed, the above mentioned 2 min photobleaching step was introduced before detection of eGFP particles in order to mimic the DOL experimental protocol. Within each cell line, the density of detected eGFP particles varied by up to one order of magnitude for individual cells. However, the average density of eGFP particles and its distribution was well comparable for the two cell lines. There was no significant difference between the eGFP densities in the two cell lines ($p = 0.35$, tested with a Wilcoxon ranksum test). In Huh7.5-pBABE-gSEP cells, an average density of 0.99 particles/ μm^2 was found and 90 % of densities lay within a range of 0.55–1.57 particles/ μm^2 . The average density of eGFP particles in Huh7.5-pBABE-LynG cells was 1.03 particles/ μm^2 while 90 % of densities lay within a range of 0.51–1.64 particles/ μm^2 . This indicates that there is no relevant difference in expression level between the gSEP construct and the LynG control construct.

As a summary, the performed bleaching experiments provided further evidence for FRET between eGFP and Halo-TMR withing the gSEP construct. However, a 2 min photobleaching step with 8.1 mW at 561 nm thoroughly photobleached Halo-TMR but did not prebleach eGFP and could thus be introduced into the experimental workflow for the determination of DOL in order to avoid FRET between the two fluorophores.

2.4. Detection of Single Particles

For the detection of single particles I chose the Point Source Detection algorithm provided by the Matlab-based tracking software u-track [36]. In short, this algorithm fits the point spread function of a single emitter with a two-dimension Gaussian intensity distribution and compares the intensity of the so detected particle against background to evaluate its significance.

In order to find the best detection parameters, I created simulated single-emitter images with different emitters densities. For this I used the ISBI Challenge Track

Generator [38] provided by the bioimage analysis software Icy [79]. Developed to simulate single-molecule diffusion trajectories for different diffusion modes and particle densities, the ISBI Challenge Track Generator can also be used to create images of immobile emitters. To mimic the cell samples, the ISBI Challenge Track Generator was provided with the pixel size of $104 \text{ nm}/\text{pixel}$ and a images size of 358×358 pixel, which corresponds to the average area of segmented Huh7.5-pBABE-gSEP cells. A total of 14 different images were generated to cover a particle density range of $0.2\text{--}2.6$ particles/ μm^2 . For each image, Gaussian-distributed noise with a standard deviation of 10 at a image depth of 8 bit of was added using Fiji. Detected particles were then colocalized with the ground truth. For the colocalization it is important to note that even for simulated data, it cannot be expected that particles are detected at exactly the ground truth coordinates. Hence, some spatial tolerance is needed that allows a particle to have a certain distance to the ground truth and still be classified as colocalized.

As an early observation, I found that the Point Source Detection algorithm detected quite a number of extremely dim particles in the noisy simulated data as well as in Huh7.5-pBABE-gSEP cell samples (Fig 2.6 left). Although these particles showed a slightly higher intensity than background (and were therefore correctly identified by the Point Source Detection algorithm) their absolute intensity was considerably lower than the typical intensity of single fluorophores. In order to adapt the sensitivity of the Point Source Detection algorithm to the expected brightness of single fluorophores, I implemented a dynamic intensity threshold that discards the dimmest population of candidate particles (see the last paragraph of section 6.6 on page 118 of material methods for details). This modification reduced the number of false-positive detections while preserving true-positive particles (Fig 2.6 right).

In a next step, I used a simulated image with a particle density of 2.0 particles/ μm^2 to optimize the particle detection parameters of the Point Source detection algorithm as well as to determine an optimal spatial tolerance value to classify colocalizations. As target parameter, the relative colocalization between simulated and detected particles was determined as the number of colocalized particles divided by the number of simulated particles. Mathematically, the relative colocalization equals the degree of labeling (see equation 2.1 on site 24). In addition, the fraction of false positive particles was calculated as the number of detected but non-colocalized particles divided by the number of detected particles. Since those target parameters require both particle detection as well as particle colocalization, the particle detection and colocalization parameters were optimized recursively.

The best empirically determined parameter set for particle detection resulted in a relative colocalization of 0.75 while the fraction of false positive detections was as low as 0.04 . The parameter set included a point spread function sigma

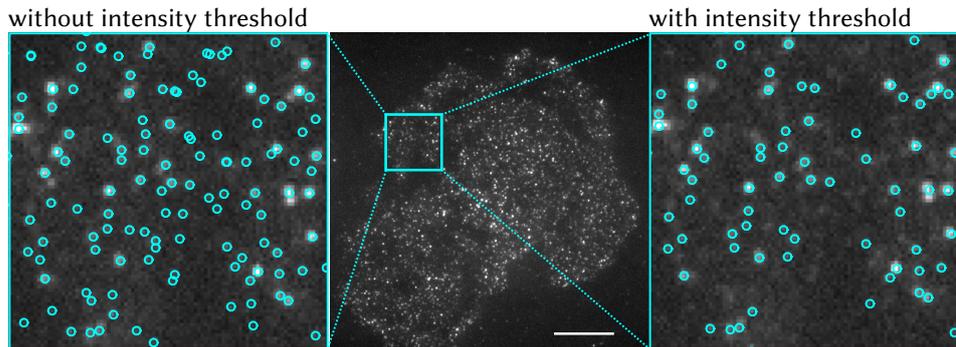


Figure 2.6.: Improved performance of u-tracks's Point Source Detection algorithm by an additional dynamic intensity threshold **a)** Without an intensity threshold, the Point Source Detection algorithm detected particles that are so dim that they were very unlikely to represent fluorophores. **b)** An additional dynamic intensity threshold that discards the dimmest population of candidate particles reduced the number of false-positives while preserving true-positive detections. Scale bar is 10 μm .

of 0.8 pixel, a maximal fitting window of 3.2 pixel and a significance level for background comparison of $\alpha = 0.005$. While isolated particles were detected with very high reliability (Fig 2.7 (a) left, overlapping point spread functions decreased the detection performance (Fig 2.7 (a) right, insufficient detection marked with a light blue arrow). In general, the fraction of particles that could be recovered decreased with increasing particle density (Fig 2.7 (b)). While for a particle density of 1.0 particles/ μm^2 still 89 % of existing particles could be detected, this ratio decreases to 78 % for a density of 2.0 particles/ μm^2 . Important to mention, the chosen parameter set did not allow for fitting point spread functions with a mixed Gaussian intensity distribution, as could be expected to occur for high emitter densities and hence overlapping point spread functions. Allowing mixture Gaussian fitting for a simulated image with 2.0 particles/ μm^2 , increased the relative colocalization only slightly to 0.76, while the fraction of false positive particles increased by 0.003. For real cell data, the intensity contributions from autofluorescence, unspecific labeling or dirt would very likely result in a even higher fraction of false positive particles as arbitrary intensity profiles can be fitted more easily with a mixture of Gaussians than with a single Gaussian. Hence, I decided that the rather minor increase in the relative colocalization does not outweigh the higher susceptibility for false positive detections.

In order to find the optimal spatial threshold for the colocalization analysis, the

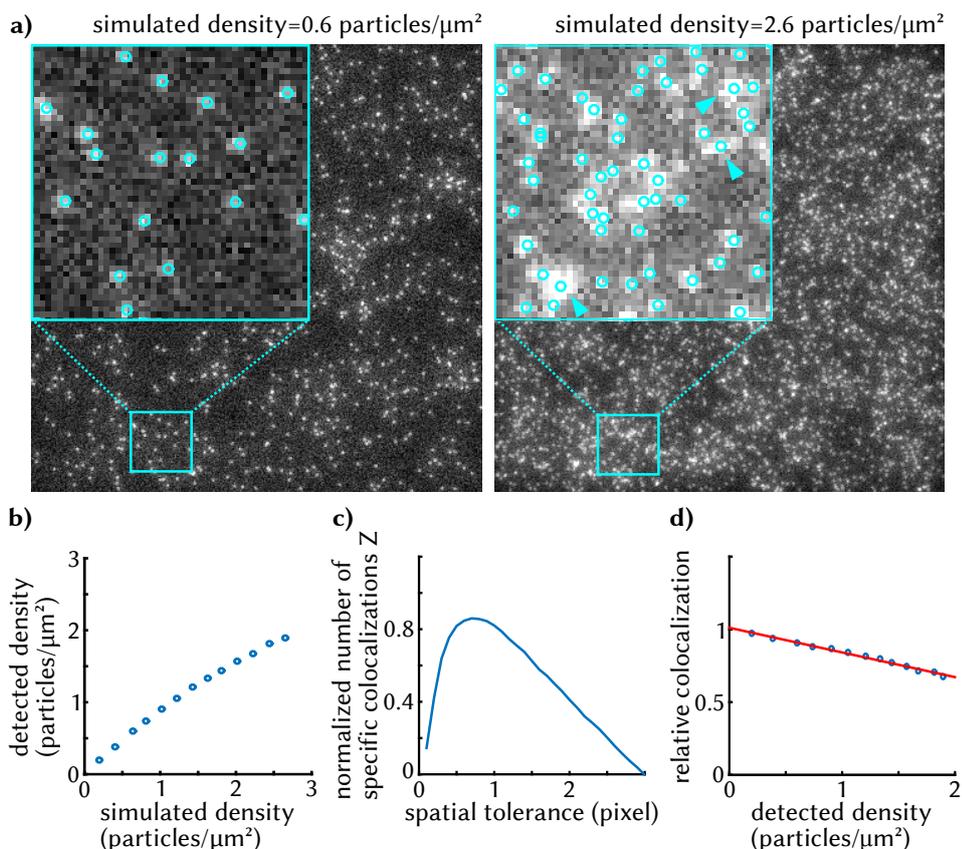


Figure 2.7.: Benchmarking the Point Source Detection algorithm based on simulated data. Images of single emitters were simulated using the bioimage analysis software Icy. **a)** U-track's Point Source Detection algorithm easily detected isolated particles, while particle detection within overlapping point spread functions in high density situation was less reliable. Examples of largely overlapping point spread functions that led to an insufficient detection are marked with light blue arrows. **b)** The reduced detection performance at high particle densities generally led to an underestimation of the particle density that increases with increasing simulated densities. **c)** To optimize the spatial tolerance of the colocalization analysis, random colocalization samples were generated by rotating the ground truth coordinates 90° clockwise. The normalized number of specific colocalizations showed an maximum at a spatial tolerance 0.8 pixel (equivalent to 83 nm), indicating that detected colocalizations are mostly specific up to this threshold. **d)** The relative colocalization between simulated and detected particles was found to decrease with increasing particle density. This relative colocalization represents the maximal relative colocalization that can be expected for a given particle density. Hence, particle density limits the dynamic range of the degree of labeling analysis.

relative colocalization of detected particles with the ground truth was determined for different spatial tolerances ranging from 0.1–3 pixel. As a control for random colocalization, detected particles were colocalized with the ground truth coordinates, however rotated 90° clockwise. For each spatial tolerances t , a normalized number of specific colocalizations Z was calculated by normalizing the number of colocalized particles between detected particles and unrotated ground truth N_{match} and between detected particles and random samples N_{random} to their respective maximal value and subtracting these normalized values from each other (Fig 2.6 (b)).

$$Z(t) = \frac{N_{match}(t)}{\max(N_{match})} - \frac{N_{random}(t)}{\max(N_{random})} \quad (2.2)$$

The normalized number of specific colocalizations showed a distinct maximum at a spatial tolerance of 0.8 pixel (Fig 2.7 (c), equivalent to 83 nm). This means that up to this threshold, with every increase of the spatial tolerance more colocalized particles are detected for the unrotated samples than for the rotated, random samples, indicating that these colocalizations are mostly specific. Above this threshold, the number of colocalizations increases more for the random sample than for the unrotated sample, indicating that those colocalizations are likely to be random. Hence, an optimal spatial tolerance of 0.8 pixel was derived from the simulated single-molecule images. This spatial tolerance was consistent for ground truth particle densities above 1 particle/ μm^2 . For lower particle densities, it decreased slightly to 0.7 pixel for a particle density of 1 particle/ μm^2 and to 0.6 pixel for particle densities below 1 particle/ μm^2 .

As high particle densities generally hinder particle detection, it may also reduce the colocalization between simulated and detected particles. Indeed, while for very low particle densities, almost all particles could be detected and colocalized with the ground truth, the relative colocalization was found to decrease linearly with increasing particle densities (Fig 2.7 (d) and equation 2.3, please note that this relation uses the detected particle density and not the ground truth density). This indicates that the maximal relative colocalization is limited to a value L that depends on the detected particle density P . Knowing this relationship however, allows to adjust any measured relative colocalization $\Theta_{measured}$ to this limited dynamic range as long as the particle density P is known. Hence, this linear relation can be used to compensate for missed particles due to a high particle density for the determination of the DOL. The linear regression in Figure 2.7 (d) yielded the following equation to correct for particle densities:

$$L = -0.17P + 1$$

$$\Theta_{true} = \frac{\Theta_{measured}}{L} \quad (2.3)$$

To sum up this section, simulating images of single emitters allowed me to optimize u-tracks's Point Source Detection algorithm towards the image quality and particle densities expected for stained Huh7.5-pBABE-gSEP cells. Furthermore I could characterize the dynamic range of the colocalization analysis in dependence on the observed particle density and obtain a linear relation to correct for this. Finally I could estimate the spatial tolerance required for the classification of colocalization, which at the same time represents the spatial error of the particle detection algorithm.

2.5. Image Registration of the eGFP, TMR and SiR Channels

The determination of the degree of labeling (DOL) of HaloTag and SNAPf-tag performed in this study is based on a three-color labeling of the gSEP construct and a colocalization analysis of particles in the three spectral channels (eGFP, Halo-TMR and SNAP-SiR in the blue, green and red channel, respectively). In the experimental setup, all three fluorophores are imaged consecutively in the same field-of-view from long to short excitation wavelengths, i.e. from red to blue, with an additional photobleaching step after imaging the TMR channel as described in section 2.3.1 on page 32. Although all three channels share the same emission path to the camera, the three resulting images are unlikely to perfectly overlay due to wavelength-specific optical aberrations and the manual change of fluorophore-specific clean-up filters in the emission pathway. To correct for this, I implemented a fully automated linear image-registration routine that aligns the three spectral channels with a precision in the one pixel range. In the following section I will present the development and validation of this image-registration routine as well as its performance based on the experimental data.

2.5.1. Development of a Linear Image Registration Based on Fiducial Markers

To create a set of training data for image registration, multicolor beads (TetraSpeck Microspheres, 0.2 μm , Thermo Fisher Scientific) that are clearly visible in all three spectral channels were deposited as fiducial markers on the glass surface of a clean

LabTek and covered with a reducing-oxidizing (ROXS) buffer (see section 6.7.1 on page 119 of material and methods). For different field-of-views, short movies of 50 frames with an exposure time of 50 ms were recorded in all three spectral channels (excitation at 480 nm, 561 nm and 640 nm correspond to the blue, green and red spectral channel, respectively). Emission in the blue and in the green were cleaned up using a 525/50 and a 605/70 bandpass filter, respectively, while no bandpass filter was used for emission in the red. This was to mimic the setup that was used to image the stained Huh7.5-pBABE-gSEP cells as closely as possible.

For each channel, an average intensity projection image of the first 10 frames was used as input for the particle detection step (Fig 2.8 (a)). As long as the observed particles are immobilized, generating a temporal projection is an effective way to reduce noise and eliminate single-molecule blinking. Particle detection was performed using the modified Point Source Detection algorithm provided by the Matlab-based tracking package u-track. As described in section 2.3.1, a dynamic intensity threshold was added to the algorithm to reduce the number of false-positive detection. Based on the coordinates of detected particles in all three spectral channels, a scaling step and a translation step was applied to project detected green and red particles onto the respective blue particles. This means, the blue channel was used as the reference channel to which the other two channels were aligned.

As a first step, the assignment of particle identities in the three different channels was performed using a nearest neighbour algorithm. Assuming that the three channels were already broadly pre-aligned, each green and red particle was assigned to the blue particle with the smallest Euclidean distance (nearest neighbour). Since the number of detected particles can be different for each channel (not so much for multicolor beads but certainly for cell images), particle assignment was performed based on the channel with fewer detected particles to avoid false-assignments due to missing particle partners. As a quality measure for the image registration, the distance distribution between all nearest neighbour particle pairs was determined after each step of the image registration for the whole image (Fig 2.8 (b), bottom).

When the three spectral channels were overlaid without any image registration, a small offset between the channels was visible (Fig 2.8 (b), left column). This offset was different in orientation and magnitude for different regions of the image. In general, the offset scaled with the spectral distance, i.e it was larger between the red and the blue channel than between the green and the blue channel. This was also reflected in the distribution of nearest neighbour distances. Without image registration, the distance between nearest neighbours varied approximately from zero to two pixels (which equals around 208 nm). The center of the nearest neighbour distance distribution, i.e. the average offset, was different for the two image dimensions and by tendency larger between the red and the blue channel than

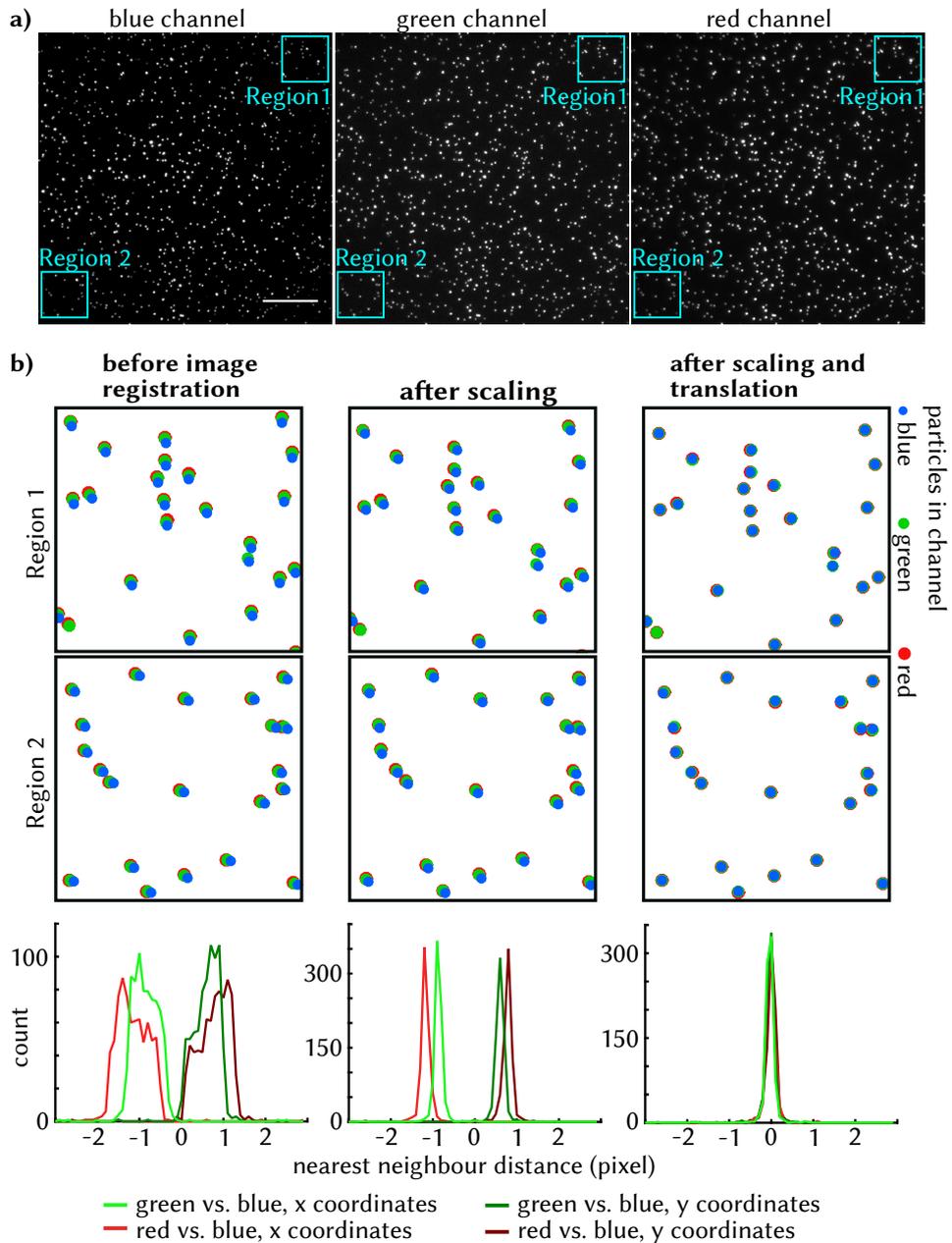


Figure 2.8.: Principles of the image registration algorithm. a) Multicolor beads were used as a training set. These beads are visible in all three spectral channels. Scale bar is $10\ \mu\text{m}$ b) The registration included a scaling step and a translation step to align detected particles in the green channel and the red channel with their respective next neighbours in the blue channel. Before image registration, the three channels showed a slight offset that differs in magnitude and orientation for different image regions. This was reflected in a broad and non-zero distribution of next neighbour distances (left). After the scaling step, the average offset was still present, however it was consistent over the image (center). This is reflected by a sharp next neighbour distance distribution that peaks at the absolute translational offset. After subtracting the translational offset, the three channels were aligned with remaining next neighbour distances below one pixel (right).

between the green and the blue channel. Without the use of image registration, a spatial tolerance of at least two pixels would be required in order to assign all particles to their respective next neighbours. While this may be well acceptable in the case of sparsely deposited beads, it may lead to conflicting and false assignments for high density labeling in cells.

Therefore, the first step of the image registration was a centered scaling of particle coordinates. For this, a scaling factor was subtracted from each coordinate of the particle in proportion to the distance of the particle to the center of the image. In other words, particles in the very center of the image were left unscaled, while particles at the image edges were corrected with the full value of the scaling factor (equation 2.4).

$$\begin{aligned} X_{i,scaled} &= X_i - \frac{X_i - 256}{256} * S_{i,X} \\ Y_{i,scaled} &= Y_i - \frac{Y_i - 256}{256} * S_{i,Y} \end{aligned} \quad (2.4)$$

,where X_i and Y_i are the x-coordinate and y-coordinate of particle i and $S_{i,X}$ and $S_{i,Y}$ are the respective scaling factors. This scaling step was performed for both dimensions of the red and the green channel individually.

The optimal scaling factors were identified by minimizing the standard deviation of the resulting nearest neighbour distance distribution. For this, the scaling factor was varied in a range of -2 to 2 in discrete steps of 0.05 and the standard deviation of the resulting next neighbour distribution was calculated for each scaling factor. The scaling factor that resulted in the smallest standard deviation was chosen for registration. After this scaling step, all nearest neighbours showed roughly the same distance, while the average offset between the spectral channels was still preserved (Fig 2.8 (b), center column). This translational offset was subtracted in the final step of the registration, resulting in well aligned channels with nearest neighbour distances well below one pixel and an average distance of zero (Fig 2.8 (b), right column).

2.5.2. Consistency of Fiducial Marker Registration Results

Throughout imaging of the stained and fixed Huh7.5-pBABA-gSEP cells to determine the DOL of HaloTag and SNAPf-tag, multicolor beads samples were regularly recorded with the aim to use the registration parameters from these easy-to-register beads sample to register the presumably not-so-easy-to-register cell images. To justify the transfer of registration parameters between samples, it had to be shown that registration parameters are consistent for a set of measurements, i.e. the scaling and translation between the channels does not considerably change from

one sample to the next.

To evaluate their consistency, the registration parameters of all beads samples were compared (Fig 2.9 (a)-(b)). In total, a set of 22 different beads samples were recorded over three days. The mean translation between the green and the blue channel was -0.36 pixel in x and -0.39 pixel in y . Between the red and the blue channel, the mean translation was -0.96 pixel in x and -0.66 pixel in y . However, the translation values from different samples differed considerably (Fig 2.9 (a)) and could not be grouped into measuring sets or measuring days. To align the green and blue channels, translations from -1.46 pixel to 1.19 pixel in x and from -0.84 pixel to 0.37 pixel in y were necessary. When aligning the red channel with the blue channel, translations from -3.27 pixel to 1.63 pixel in x and from -1.31 pixel to 0.34 pixel in y were obtained. Due to these large variations between individual samples, a transfer of beads-derived translation parameters to cell samples had to be rejected. Interestingly however, a significant linear correlation could be found between the translation of the red channel and the translation of the green channel (Fig 2.9 (c)). For the translation in x direction, a Pearson correlation coefficient of $R = 0.94$ and p -value of $p = 1.6e - 10$ was found. In y -direction, the Pearson correlation coefficient was $R = 0.88$ with a p -value of $p = 5.9e - 08$. This indicates that the orientation and the magnitude of the translational offsets of the green channel and of the red channel towards the blue channel are related within one sample and can in principle be estimated from each other.

In contrast to the translation, the scaling factor showed considerably less variability between different samples (Fig 2.9 (b)). Except for single outliers, the derived scaling factors cluster within an interval of around 0.5 pixel in both dimensions and for both channels. The mean scaling factor for the green channel was 0.55 in x (± 0.22) and 0.49 in y (± 0.09). The mean scaling factor for the red channel was 0.67 in x (± 0.18) and 0.57 in y (± 0.17).

Using the mean translation and mean scaling factor to register all 22 beads samples resulted in nearest neighbour distances that were distributed around zero but showed distinct populations for those samples where the optimal translation was clearly different from the mean value (Fig 2.9 (d), left). In this case, a spatial tolerance of around 3 pixel (equivalent of 312 nm) for the red channel and around 1.5 pixel (equivalent of 156 nm) for the green channel would be necessary for a correct particle assignment. Hence for dense particles, registration with mean registration parameters may not be suitable.

To test the accuracy of the correlation between the translations of the green and the red channel, the translation of the green channel was individual obtained for all 22 samples and the corresponding translation for the red channel was derived from the correlation depicted in Figure 2.9 (c). For both channels, a mean scaling factor was used. The resulting nearest neighbour distance distribution was sharply

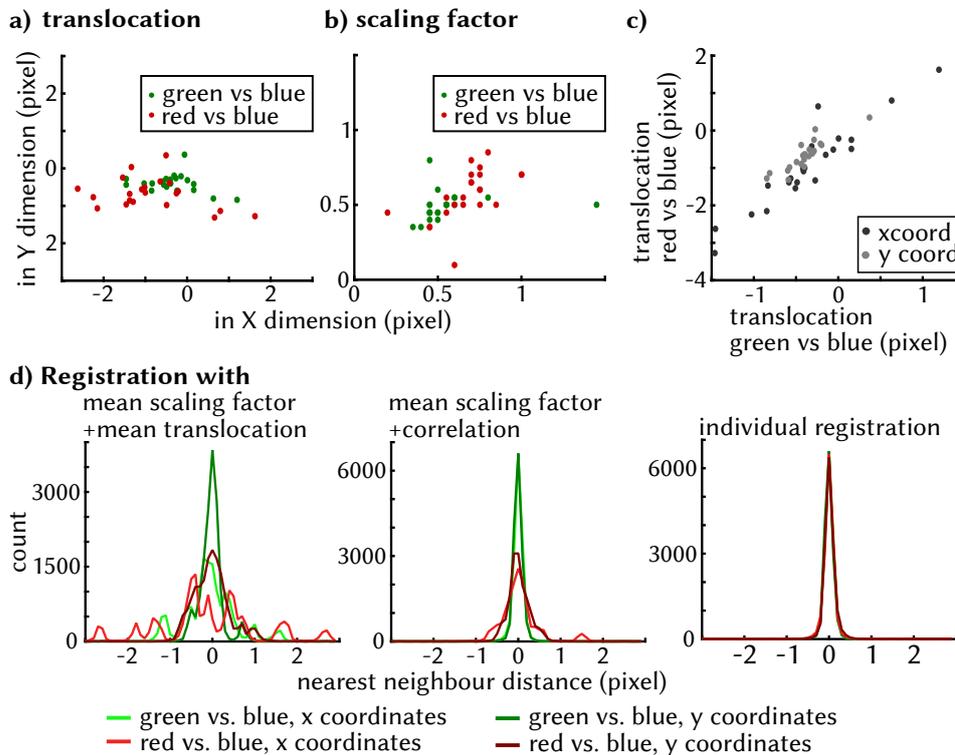


Figure 2.9.: Variability of registration parameters for different fiducial marker samples. Registration parameters were derived for a total of 22 multicolor beads samples over three days. **a)** The translation parameters obtained from different samples differed by several pixel for both channels and both dimensions. Hence the translation from one sample cannot be transferred to a different sample. **b)** The corresponding scaling factors showed less variability and allowed to use a mean scaling factor for all samples. **c)** The translation of the green channel correlated significantly with the translation of the red channel. **d)** When only mean registration parameters were used for registration, the nearest neighbour distance distribution of all 22 samples showed distinct populations with a non-zero distance (left). Using a mean scaling factor for both channels and obtaining the translation for the red channel via the correlation to the green channel largely reduced the variability of nearest neighbour distances in the red channel while the respective distribution in the green channel narrowed around zero (center). When both channels were individually registered, the resulting nearest neighbour distances were below 0.5 pixel (right).

centered at zero the green channel (Fig 2.9 (d), center) with a width comparable to the full individual registration (compare Fig 2.9 (d), right). This indicates that using a mean scaling factor for all samples of one channel did not considerably worsen the registration precision. The nearest neighbour distances of the red channel showed considerably less variability when compared to a registration with mean values. Especially the number of sub-populations with a clear non-zero distance could be nearly eliminated. In this case, a spatial tolerance of 1 pixel would be enough to correctly assign corresponding particles between the red channel and the blue channel, while the necessary spatial tolerance between the green channel and the blue channel was below 0.5 pixel. This indicates, that the correlation between the translation of the green and red channel could help to determine non-accessible translation parameters.

Finally, individual registration of all 22 samples resulted in nearest neighbour distances below 0.5 pixel (which represents 52 nm) for both channels and both dimensions (Fig 2.9 (d), right).

As a summary of this section, I found that the registration variability between different samples is too large to transfer registration parameters from beads samples to cell samples or to use mean registration values. Therefore, individual registration of the cell samples was necessary. If the registration parameters can only be determined for one channel, the correlation between the translation of the green and the red channel in principle allows to indirectly estimate the translation of one channel from the other.

2.5.3. Registration of Cell Images

After the successful validation of an automated registration routine based on fiducial marker samples and the conclusion that registration parameters cannot be transferred between samples, I tested the registration towards its use to align the eGFP, Halo-TMR and SNAP-SiR channels of individual stained Huh7.5-pBABE-gSEP cell images. For this, Huh7.5-pBABE-gSEP and Huh7.5-pBABE-LynG cells were stained and fixed as described in section 6.7.3 on page 119. In total, 40 different labeling conditions were used for each cell types and 10–12 replicates per labeling condition were recorded (see also Fig 2.11 on page 49). For each field of view, a 100 frame movie with an exposure time of 50 ms was recorded for each spectral channel. In order to reduce noise and remove unfixed, mobile particles, an average intensity projection image of the first 10 frames was created and used as input for particle detection. In Huh7.5-pBABE-LynG cells, a high amount of unspecific but fast decaying background fluorescence was detected in the red (SiR) channel. Therefore, an average intensity projection image of frames 11–20 was created for these samples.

For all cells, the area of the cell was detected via segmentation of the eGFP channel. As mentioned in section 2.3.1, the segmentation step included a rolling ball algorithm to extract local average intensities followed by a low-pass filter to eliminate features smaller than 25 pixel, a median filter with a radius of 30 pixel to homogenize the intensity throughout the cell and finally a threshold based on Fiji's Triangle algorithm. For the segmented area, particle detection was performed using the modified Point Source Detection algorithm described in section 2.4 starting on page 35.

Based on the obtained coordinates, a fully automated registration as described in subsection 2.5.1 on page 40 was performed in order to align the Halo-TMR channel and the SNAP-SiR channel with the eGFP channel, which served as the reference channel. As a reminder, the registration algorithm includes a scaling step and a translation step. Since the scaling factors were found to be relatively consistent based on fiducial marker samples, all cell samples were scaled using the average scaling factor obtained from these samples. These were 0.55 and 0.49 for the x- and y-coordinates of Halo-TMR particles and 0.68 and 0.57 for the x- and y-coordinates of SNAP-SiR particles, respectively.

The translation offset was determined for every cell sample individually, if possible. As described in subsection 2.5.1, the offset is found as the most abundant distance in the nearest neighbour distance distribution between the particles of two channels under the assumption that the two channels are broadly pre-aligned. For multicolor beads, which are clearly visible in both channels, the translation offset is a very prominent peak (see Fig 2.10 (a) or Fig 2.8 (b) on page 42). This is because the offset is way smaller than the distance between particles and all particles have a corresponding partner in the other channel. Therefore, false assignments are very unlikely. True corresponding particle pairs will hence show the smallest Euclidean distance and contribute to the nearest neighbour distance distribution.

For cell samples however, it gets more complicated due to the higher particle density and the unknown degree of labeling. In the case of the gSEP construct, an insufficient labeling of the protein tags will result in particles that are visible in only one of the channels (e.g. only in the eGFP channel when HaloTag was not labeled). Since in this case no corresponding particle can be found in the other channel, the nearest neighbour distance will be the distance to a close but unrelated particle and will therefore not represent the translation offset. However, if there is a large enough population of particles that are visible in both channels, then this translation offset will repeatedly be observed compared to the more random distances between unrelated particles.

In fact, the nearest-neighbour distance distribution of stained Huh7.5-pBABE-gSEP cells usually showed distinct peaks that can be assigned to the translational offset. Figure 2.10 (b) shows such a peak for nearest neighbours between the

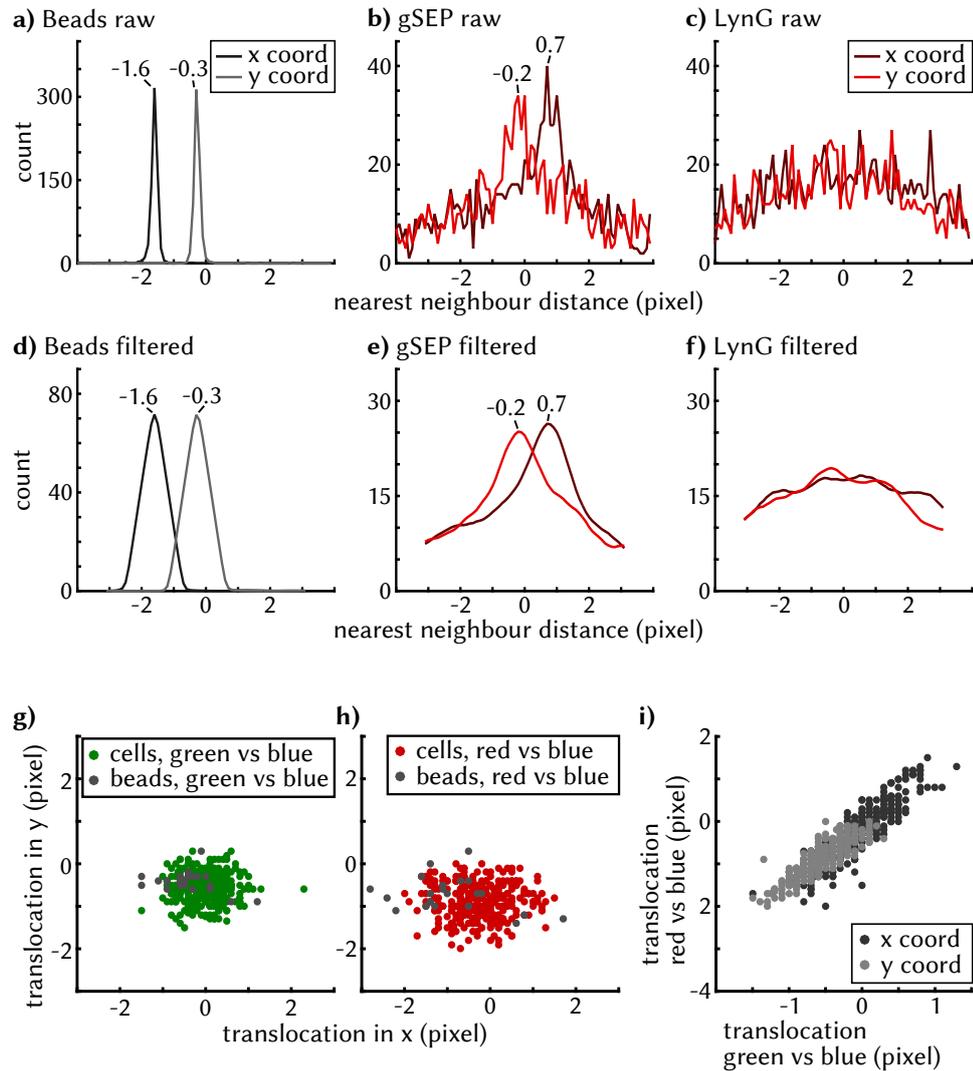


Figure 2.10.: Automated registration of cell images. The samples in a)-f) show nearest neighbour distance distributions between particles of the eGFP (blue) channel and the SNAP-SiR (red) channel. **a)** For multi-color beads as fiducial markers, the translation offset between the two channels could be found as a prominent peak in the nearest-neighbour distance distribution. **b)** For a Huh7.5-pBABE-gSEP cell, which was stained with 5 nM SiR-BG overnight, this peak was less prominent but detectable for both coordinates. **c)** A Huh7.5-pBABE-LynG cell that was stained with 100 nM SiR-BG overnight did not show a specific peak in the nearest neighbour distance distribution even though the number of detected nearest neighbours was comparable to the stained Huh7.5-pBABE-gSEP cell. **d)-f)** Two consecutive sliding average filters with a window size of 0.9 pixel conserved the localization of the translation offset in beads samples and facilitated its detection for stained cell samples. **g)-h)** Compared to the beads samples, the translation offsets derived from cell samples cluster slightly closer around 0. The registration of the SNAP-SiR channel (red) required by trend larger translations than the registration of the Halo-TMR channel (green). **i)** The translational offset of the Halo-TMR channel correlated significantly with the translational offset of the SNAP-SiR channel.

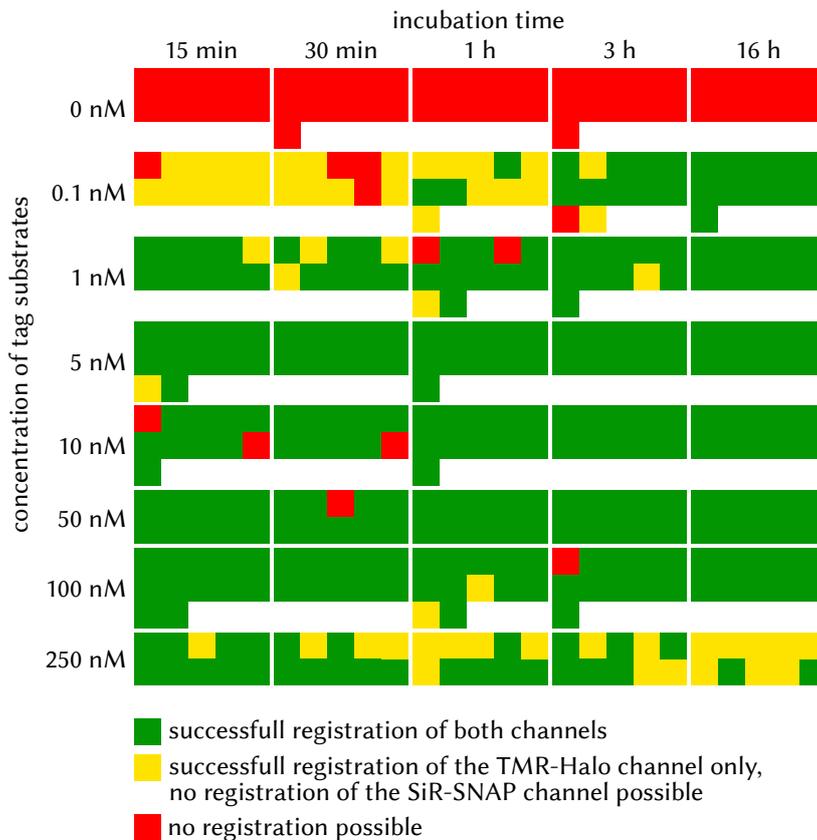


Figure 2.11.: Success of automated registration for different labeling conditions. Huh7.5-pBABE-gSEP cells were labeled under varying incubation times and concentrations of SiR-BG and HaloTag TMR ligand. The registration success is color coded for every replicate of every condition. A replicate where both the Halo-TMR channel as well as the SNAP-SiR channel could be registered to the eGFP channel is represented by a green square. Yellow squares identify replicates where only the Halo-TMR channel could be registered. If none of the channels could be registered to the eGFP channel, the replicate is represented as a red square. A successful registration of the Halo-TMR channel was generally possible for all tested conditions. Registration of the SNAP-SiR channel in contrast was impeded by either low numbers of specific SNAP-SiR particles (in the case of low concentration and short incubation times) or by a high number of unspecific particles (in the case of high concentration and long incubation times). All labeling conditions were also applied to Huh7.5-pBABE-LynG cells with a comparable number of replicates. For these cells, no registration was possible.

eGFP channel and the SNAP-SiR channel for a Huh7.5-pBABE-gSEP cell that was stained overnight with 5 nM SiR-BG. For comparison, Figure 2.10 (c) shows the same distribution for a Huh7.5-pBABE-LynG cell that was stained overnight with 100 nM SiR-BG. Since the LynG construct does not contain the SNAPf-tag, particles in the SNAP-SiR channel arise from unspecific labeling and are therefore unrelated to particles in the eGFP channel. Although the number of detected particles, and therefore the number of nearest neighbour distances, was roughly comparable to the stained Huh7.5-pBABE-gSEP cell, no prominent distance peak could be detected.

In order to extract the translational offset, I used two consecutive sliding average filters each with a window size of 0.9 pixel to smooth the nearest neighbour distance distribution (Fig 2.10 (d)-(f)). For the beads sample, the very sharp peak broadened due to the averaging, however the peak positions on the x-axis were conserved (Fig 2.10 (d)). In case of the cell samples, smoothing the nearest neighbour distance distribution allowed for a better localization of the peak as well as for an evaluation of its quality. For all samples, maxima of the filtered nearest neighbour distance distribution were used as translation offset if their absolute count was above 6 and if this count was at least 1.4 times higher than the average count within an interval of -3 to 3 pixels along the x-axis. This quality criterion was empirically determined.

With regard to the registration of the Halo-TMR channel with the eGFP channel, the quality criterion was met by 358 out of 367 images of stained Huh7.5-pBABE-gSEP cells (Fig 2.11). 301 of these images also met the quality criterion for registration of the SNAP-SiR channel with the eGFP channel. As a negative control, none of the 52 unstained Huh7.5-pBABE-gSEP cells and none of the 406 similarly stained Huh7.5-pBABE-LynG cells showed a peak in the nearest neighbour distance distribution that met the quality criterion.

As a trend, HaloTag labeling produced enough corresponding particle pairs to allow for an automated registration for almost all tested labeling conditions. Only for very low HaloTag TMR ligand concentrations of 0.1 nM combined with incubation times below 1 h, registration failed in up to 3 out of 10 samples, probably due to insufficient particle numbers. This problem was even more prominent for SNAPf-tag labeling, where hardly any successful registration could be obtained for a concentration of 0.1 nM SiR-BG unless cells were incubated for at least 3 h. In addition, registration of the SNAP-SiR channel was often not possible for high SiR-BG concentrations combined with long incubation times. For these conditions, a high degree of unspecific labeling could be detected, that likely hampered the detection of specific particles. The amount of unspecific staining is described in detail in section 2.6.1 on page 62.

Based on all successful registrations, the average translational offset between

the Halo-TMR channel and the eGFP channel was found to be -0.01 pixel in x and -0.57 pixel in y (Fig 2.10 (g)). Between the SNAP-SiR channel and the eGFP channel, the average translation offset was -0.24 pixel in x and -0.88 pixel in y (Fig 2.10 (h)). These mean translational offsets are slightly different from their respective mean values derived from beads samples in section 2.5.2. In x -dimension, this difference is significant at a significance level of $\alpha = 0.001$, indicating that registration parameters may generally be influenced by the optical properties of the sample.

As was already seen for the beads samples, a significant correlation was found between the translation offsets of the Halo-TMR (green) channel and the SNAP-SiR (red) channel (Fig 2.10 (i)). For the translation in x direction, a Pearson correlation coefficient of $R = 0.89$ ($p = 7.3e - 105$) was found. In y -direction, the correlation coefficient was $R = 0.85$ ($p = 9.2e - 87$). A linear regression yielded the following relationship:

$$\begin{aligned} offset_{red,X} &= 1.41 * offset_{green,X} - 0.18 \\ offset_{red,Y} &= 1.09 * offset_{green,Y} - 0.26 \end{aligned} \quad (2.5)$$

,where $offset_{red,X}$ and $offset_{red,Y}$ are the translations needed to align the SNAP-SiR channel with the eGFP channel in x and y direction while $offset_{green,X}$ and $offset_{green,Y}$ are the respective translations needed to align the Halo-TMR channel with the eGFP channel. Using this linear relationship between the translation offsets of the Halo-TMR channel and the SNAP-SiR channel, the translation offset of the SNAP-SiR channel could be estimated for those replicates where only the Halo-TMR channel was successfully registered.

Based on this correlation and the average translational offsets of all successfully registered cells, the following registration workflow was created in order to register all samples as good as possible: In a first step, the coordinates of detected particles in the Halo-TMR and in the SNAP-SiR channel were scaled with the average scaling factors derived from all 22 beads samples as described in section 2.5.2 on page 43. Next, the translation offset between the Halo-TMR channel and the eGFP channel was derived. If this failed, both the Halo-TMR and the SNAP-SiR channel were registered based on the mean translation offset from all successfully registered cells. These replicates are represented as red squares in Fig 2.11, since their registration quality might be low. Otherwise, if the translation offset of the Halo-TMR channel could be successfully obtained, the Halo-TMR channel was registered with these individual values and an attempt to obtain the translation offset of the SNAP-SiR channel was started. If this failed, the translation offset of the SNAP-SiR channel was determined from its correlation with the translation offset of the Halo-TMR channel. These replicates are represented as yellow squares in Fig 2.11. Otherwise,

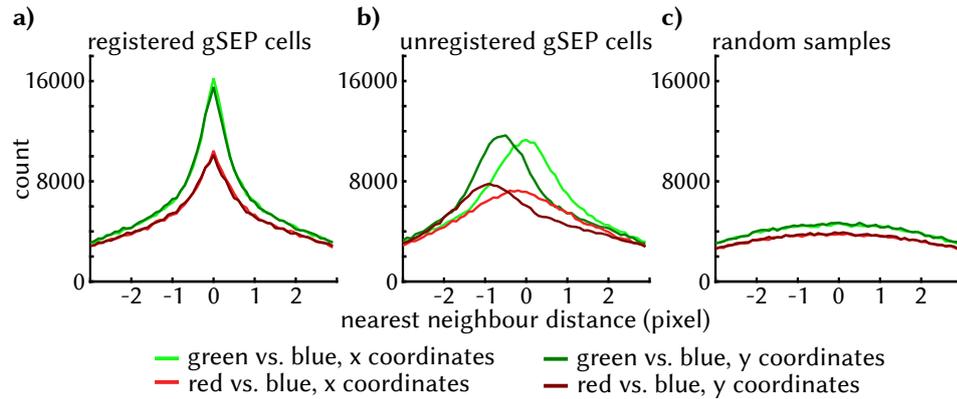


Figure 2.12.: Validation of cell image registration. All Huh7.5-pBABE-gSEP cell samples were registered using the automated registration routine described in this section. All graphs show the nearest neighbour distance distribution between particles of the eGFP channel and particles of the Halo-TMR or the SNAP-SiR channel for both dimensions. **a)** After registration, the average nearest neighbour distance of all cell samples was zero for both spectral channels and both coordinates. **b)** In contrast, unregistered cell samples showed clear non-zero maxima that corresponded to the average translational offsets depicted in Figure 2.10 (g)-(h). **c)** Random samples, generated by rotating the Halo-TMR channel and the SNAP-SiR channel 90° clockwise, did not show any prominent nearest neighbour distances.

if the translational offset could be obtained for both channels individually, the replicate is represented as green squares in Fig 2.11 and for both channels the individual translation offsets were used to align them with the eGFP channel.

Using this registration routine, the average nearest neighbour distance based on all Huh7.5-pBABE-gSEP cell samples could be centered to zero for both coordinates and both channels (Fig 2.12 (a)). For the unregistered samples in contrast, the nearest neighbour distance distribution recovered at the mean translation offsets of 0 pixel in x and -0.6 pixel in y for the Halo-TMR (green) channel and -0.3 pixel in x and -0.9 pixel in y for the SNAP-SiR (red) channel (compare Fig 2.12 (b) and Fig 2.10 (g)-(h)).

As a control for random colocalization, random nearest neighbour distance samples were created by rotating the Halo-TMR and SNAP-SiR channel of each cell sample 90° clockwise. The nearest neighbour distance distribution between particles of these random controls and particles of the corresponding eGFP channel showed no considerable peak for any of the rotated channels (Fig 2.12 (c)).

As a summary of this section, I developed a automated image registration routine that allowed to align the three spectral channels of individual Huh7.5-pBABE-gSEP cell images without the need for fiducial marker calibration samples. This is a major prerequisite for the following colocalization analysis.

2.5.4. Determination of the Spatial Colocalization Tolerance

After registration, the coordinates of Halo-TMR and SNAPf-SiR particles were compared with the coordinates of eGFP particles in order to decide if particles of the different spectral channels colocalize or not. As already mentioned in section 2.4, corresponding particles can not be expected to have exactly the same coordinates due to localization errors introduces for example during particle detection and due to a non-perfect image registration. Therefore, a spatial tolerance distance has to be determined that represents the threshold distance between two particles of different spectral channels below which the two particles are classified as being colocalized. This spatial tolerance was estimated for simulated single-molecule images in section 2.4 on page 39 during the optimization of particle detection and an analogous estimation was performed for cell samples. For cell samples, the number of colocalized particles derived from all Huh7.5-pBABE-gSEP cell images was compared to the number of colocalized particles derived from the corresponding, 90° rotated random samples for different spatial tolerances ranging from 0.1–4 pixel (which equals 10–416 nm). Following equation 2.6, for each spatial tolerance t a normalized number of specific colocalizations Z was calculated by normalizing the number of colocalized particles from gSEP cell samples N_{gSEP} and random samples N_{random} to their respective maximal value and subtracting these normalized values from each other.

$$Z(t) = \frac{N_{gSEP}(t)}{\max(N_{gSEP})} - \frac{N_{random}(t)}{\max(N_{random})} \quad (2.6)$$

This calculation was performed for the Halo-TMR channel and the SNAP-SiR channel individually but the same spatial tolerance was used for the x and y dimension. The normalized number of specific colocalizations in dependence to the spatial tolerance is shown for both channels in Fig 2.13 (a).

For the SNAP-SiR channel the normalized number of specific colocalizations showed a maximum at a spatial tolerance of 1.4 pixel (equivalent of 146 nm), while this maximum for the Halo-TMR channel was at spatial tolerance of 1.3 pixel (equivalent of 135 nm). This means that with every increase of the spatial tolerance up to these thresholds, more colocalized particles were detected in the cell samples

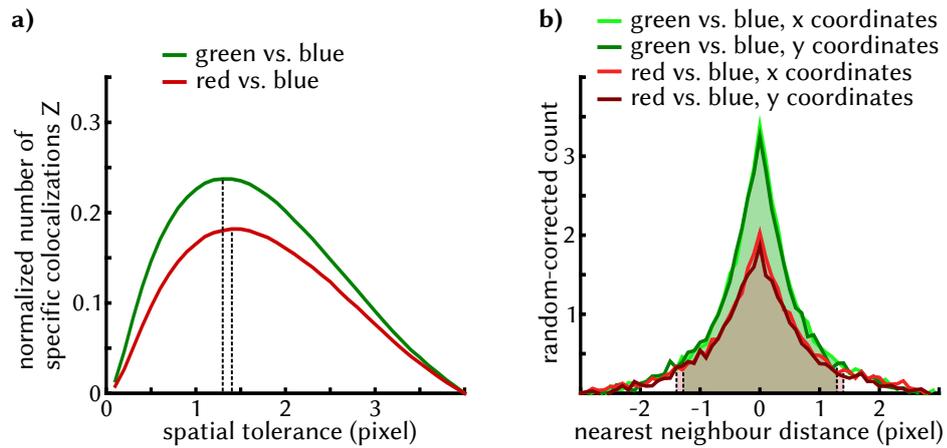


Figure 2.13.: Determination of the spatial colocalization tolerance. **a)** To estimate the spatial tolerance for the assignment of colocalized particle pairs, a normalized number of specific colocalizations was determined based on all gSEP cell samples and the respective random samples. For the TMR-Halo channel and the SNAP-SiR channel, a maximal difference could be found for a spatial tolerance of 1.3 pixel and 1.4 pixel, respectively, indicating that colocalizations are mostly specific up to this threshold. **b)** When corrected with random nearest neighbour distances, 90 % of nearest neighbour particle pairs from registered cell samples show a distance between -1.3 pixel for Halo-TMR/eGFP pairs and -1.4 pixel for SNAP-SiR/eGFP pairs (area marked in light green for Halo-TMR particles and light red for SNAP-SiR particles).

than in their respective random control samples, indicating true colocalizations of the same underlying gSEP constructs. For spatial tolerances above this threshold, the increase of colocalizations was larger for the random samples than for the cell samples, indicating that particles that were classified as colocalized above this threshold may only colocalize by chance and may not represent the same gSEP construct.

A further justification of these thresholds resulted from the random-corrected nearest neighbour distance distribution of registered Huh7.5-pBABE-gSEP cell samples (Fig 2.13 (b)). For this, the nearest neighbour distance distributions of all registered Huh7.5-pBABE-gSEP cell samples (as shown in Fig 2.12 (a)) and the nearest neighbour distance distribution of the corresponding random samples (as shown in Fig 2.12 (c)) were normalized to their respective minimum value and subtracted from each other, in order to remove random nearest neighbour

assignments. Based on this corrected next neighbour distance distribution, 90 % of nearest neighbours between SNAP-SiR particles and eGFP particles were included within a interval of 1.4 pixels and would be classified as colocalized when a spatial tolerance of 1.4 pixel was used. The same is true for nearest neighbours between Halo-TMR particles and eGFP particles where 90 % showed a distance within 1.3 pixels in both dimensions.

As a summary, I could deduce a spatial distance threshold, below which particles of different spectral channels can be safely assumed to represent the same underlying molecule in order to classify them as colocalized.

2.5.5. Prevention of Multiple Assignments

For any spatial tolerance, there is the possibility that a particle in one spectral channel can find multiple particles in the other spectral channel at a distance below this threshold. Those situation can lead to multiple assignments of particles. If the number of colocalized particles is determined as the number of nearest neighbour pairs with a distance below the spatial tolerance, multiple assignments will results in a overestimation of the number of colocalized particles.

To prevent multiple assignment events, I developed a colocalization algorithm that allows each particle of one channel to colocalize with only one particle of the other channel. More specifically, the algorithm loops over all particles that have nearest neighbours at a distance below the spatial tolerance. If more than one partner is found, the particle is assigned to the first partner in the loop while all other assignments are removed. Since this loop generates an artificial order, it was run in forward and backward direction and the maximal number of colocalizations was kept.

In order to exemplify and verify the algorithm, several motifs that contain multiple assignments are displayed in Figure 2.14 (a). Particles of two different spectral channels are depicted in red and blue and neighbouring particles (horizontal, vertical or diagonal) were set to lie within the spatial tolerance. In Motif 1 only two colocalizations are possible although the number of nearest neighbour pairs below the spatial tolerance is actually 8. In fact, when multiple assignments are prevented, the colocalization detection algorithm gave a total of 2 colocalizations, compared to 8 colocalizations without prevention of multiple assignments. For motif 2, the number of colocalizations without preventing multiple assignments was 12, while preventing multiple assignments reduced this number to the expected 4. For motif 3, a maximal number of 3 colocalizations is possible. Due to the asymmetry of this motif, the direction of looping the particles may determine which particle pairs are actually assigned to each other. In fact for this motif, the forward loop gave only 2 colocalization, while the backwards loop yielded the expected number of 3.

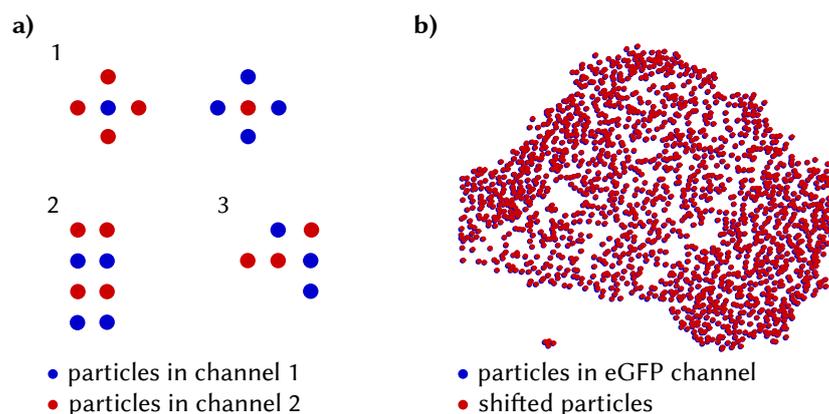


Figure 2.14.: Multiple assignments of particles. a) When a particle has multiple partner at a distance below the spatial tolerance, multiple assignment may lead to an overestimation of the number of colocalized particles. The motifs 1–3 show situations where multiple assignments between particles of different channels (blue and red) can occur. For the three motifs, the number of particle pairs below the spatial tolerance is 8, 12 and 6, respectively. Designed to prevent multiple assignments, the colocalization algorithm recovered the expected number of 2, 4 and 3 colocalizations. b) To generate a cell data-based colocalization sample with known ground truth, all 1680 particles of the eGFP channel of a Huh7.5-pBABE-gSEP cell were shifted by 1 pixel in x and y. Using on a spatial tolerance of 1.4 pixel, the colocalization algorithm obtained the expected number of 1680 colocalizations while detecting 20 multiple assignments.

Not preventing multiple assignments yielded as many as 6 colocalizations.

Because real cell data can have multiple such motifs, the colocalization detection algorithm was tested on a cell sample with known ground truth. For this, all eGFP particles of a Huh7.5-pBABE-gSEP cell image were shifted by 1 pixel in x direction and 1 pixel in y direction, in order to mimic a rather large remaining offset in a registered cell sample. The original particles as well as the slightly shifted particles are visualized in Figure 2.14 b). Using a spatial tolerance of 1.4 pixel, a total of 1680 colocalized particles between the original and the shifted channel were detected. This equals the total number of particles in the eGFP channel and since the spatial shift of 1 pixel in both dimensions is below the spatial tolerance, it is also the expected number of colocalizations. In total, 20 multiple assignments were detected, i.e. without preventing multiple assignments the number of colocalized particle would have been 1700. This indicates that for the typical spatial tolerances

and offsets found in cell images, multiple assignment of particles is not a major problem and can effectively be prevented by the colocalization detection algorithm. As a further characterization, the correction of multiple assignments was found to work also for larger offsets. When the coordinates of particles in the eGFP channel were shifted by 10 pixels, still the correct number of 1680 colocalizations could be recovered using a spatial tolerance of 11 pixel even though multiple assignments occurred in 10 385 cases.

2.6. The Degree of Labeling of SNAPf-tag and HaloTag

In section 2.5.3 on page 46, I have mentioned that I stained Huh7.5-pBABE-gSEP cells and Huh7.5-pBABE-LynG cells with HaloTag TMR ligand and SiR-BG under a total of 40 different labeling conditions, varying the concentration and incubation time of the two substrates (see Fig 2.11 for an overview of labeling conditions and section 6.7.3 on page 119 of material and methods for a detailed labeling protocol). Using the already described particle detection, image registration and colocalization algorithms, I then detected the signal of individual gSEP constructs in different spectral channels and calculated the degree of labeling (DOL) for HaloTag and SNAPf-tag following equations 2.1 on page 24 as the number of particles that colocalized in the eGFP channel and in the channel of the respective tag, divided by the total number of particles in the eGFP signal. As described in section 2.4, the maximal observable colocalization decays linearly with increasing particle density, which likewise limits the dynamic range of DOL detection. Therefore, the experimentally determined DOL was corrected for the measured particle density for every image and both protein tags individually using equation 2.3 on page 40. For all labeling conditions, the resulting DOL of HaloTag and SNAPf-tag is visualized in Figure 2.15 a) and summarized in Table 2.1.

I found that the DOL of HaloTag reached a relatively stable plateau value of 0.4–0.5 that was mostly independent of the dye concentration but required incubation times longer than 1 h (Fig 2.15 (a), left). A DOL of 0.4–0.5 means that 40–50 % of HaloTag proteins are labeled with their respective substrate. For shorter incubation time, the DOL decreases slightly to 0.3–0.4. For very short incubation times below 1 h combined with very low substrate concentration of 100 μ M, the DOL was lowest with values around 0.15. The DOL of HaloTag could not be further increased by adding a higher substrate concentration. If anything, the DOL was by tendency slightly decreased for very high substrate concentration, although the observed reduction lies well within the range of the experimental error. As will be discussed in section 2.6.1, this may be due to an increasing amount of unspecific background signal, that may hamper the detection of specific particles. However,

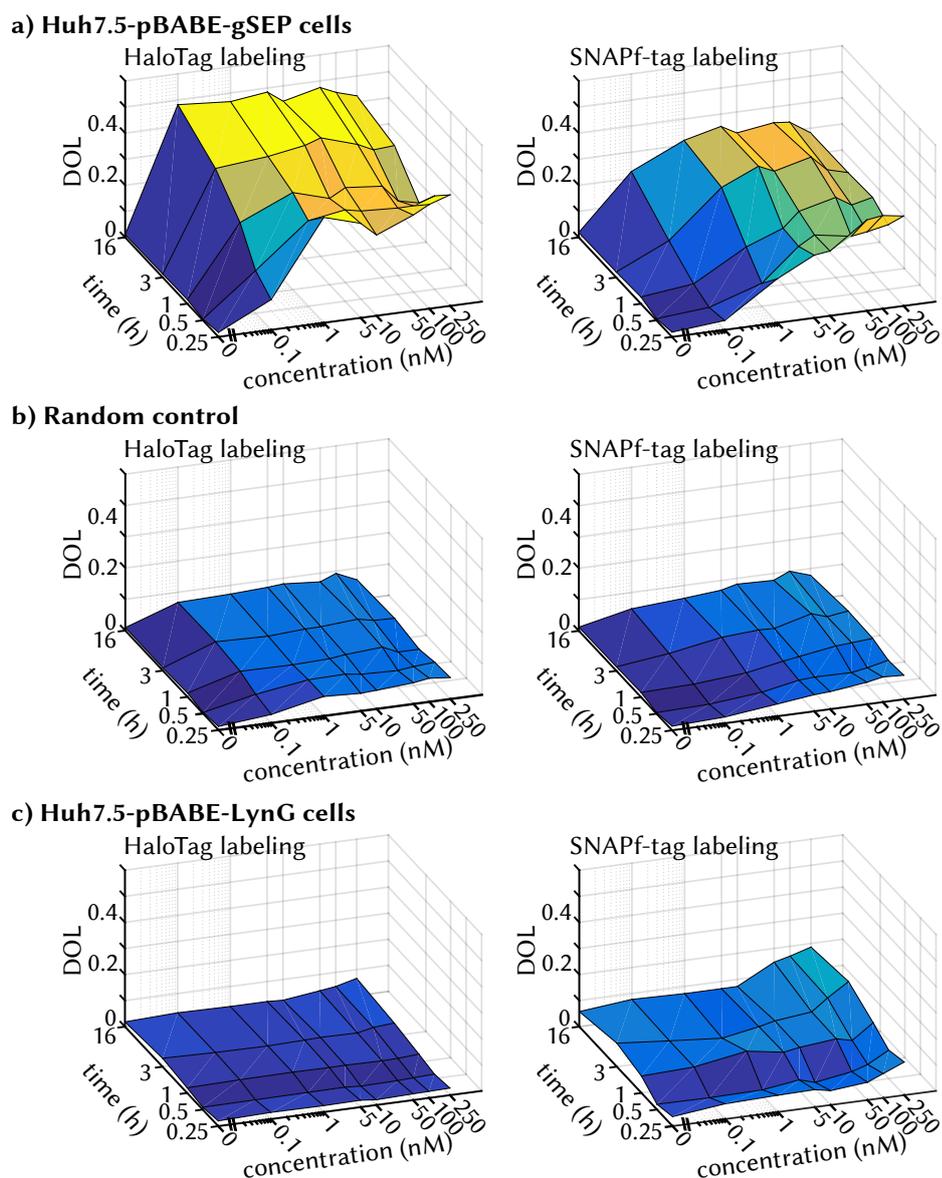


Figure 2.15.: Degree of labeling of HaloTag and SNAPf-tag. 40 different labeling conditions were tested with at least 10 replicates per condition in both Huh7.5-pBABE-gSEP cells and Huh7.5-pBABE-LynG cells.

a) In Huh7.5-pBABE-gSEP cells, HaloTag showed a DOL of 0.3–0.5 for most labeling conditions. The DOL of SNAPf-tag was significantly lower for most labeling conditions and more influenced by insufficient and unspecific labeling (Wilcoxon rank sum test with a significance level of $\alpha = 0.05$). **b)** Random control samples yielded a random colocalization below 0.1 for both tags and any labeling condition. **c)** Unspecific labeling, as measured in Huh7.5-pBABE-LynG cells, resulted in unspecific colocalization of 0.03 for HaloTag and 0.07 for SNAPf-tag, on average. Only for very high SiR-BG concentrations combined with very long incubation times, unspecific labeling results in a DOL that is approaching values measured in Huh7.5-pBABE-gSEP cells. These two control experiments indicate that the determined DOL of HaloTag and SNAPf-tag is largely independent of random and unspecific colocalization.

		Incubation time				
		15 min	30 min	1 h	3 h	16 h
Concentration of substrates	0 nM	2±0 2±1	0±0 1±0	2±0 3±0	1±0 2±1	1±0 2±1
	0.1 nM	12±4 4±1	16±4 2±1	28±5 5±2	39±5 11±4	48±7 23±6
	1 nM	40±6 17±5	33±4 9±3	37±16 16±8	41±9 29±8	47±11 32±6
	5 nM	36±7 24±6	35±6 21±6	33±8 22±7	43±2 35±5	48±10 35±4
	10 nM	31±11 25±8	34±5 25±6	36±6 27±7	45±7 33±5	44±12 32±5
	50 nM	37±6 35±7	33±6 22±6	35±9 31±6	41±8 36±3	47±6 34±4
	100 nM	40±7 36±5	33±6 23±6	29±5 27±4	38±8 31±4	45±12 33±3
	250 nM	43±4 34±3	36±6 25±4	27±6 29±4	39±5 28±3	42±10 28±3
		(DOL HaloTag ± standard deviation) * 100 (DOL SNAPf-tag ± standard deviation) * 100				

Table 2.1.: Degree of labeling of HaloTag and SNAPf-tag. Overview of the DOL of HaloTag (green) and SNAPf-tag (red) and the corresponding standard deviation for all tested labeling conditions.

even in a concentration range below 100 nM, where unspecific labeling is not an issue for HaloTag labeling, the DOL seemed largely independent of the substrate concentration while incubation time on the other hand, had a stronger influence. For incubation times longer than 3 h, the plateau value can be reached using as little as 100 pM HaloTag TMR ligand. In fact, the maximal DOL of just below 0.5 could be obtained for overnight staining with 100 pM HaloTag TMR ligand. The standard deviations of the DOL of HaloTag in Huh7.5-pBABE-gSEP cells averaged 0.06, ranging from 0.003 to 0.16.

The DOL of SNAPf-tag was significantly lower than the DOL of HaloTag for most labeling conditions (Fig 2.15 (a), right). It showed a less robust plateau value

between 0.3–0.4 that required an incubation time of at least 3 h and a substrate concentration of 5–100 nM. Compared to HaloTag labeling, SNAPf-tag labeling seemed to be more prone to insufficient labeling when concentrations below 5 nM or incubation times shorter than 3 h were used. At the same time, SiR-BG concentrations above 100 nM likewise resulted in a by-tendency lowered DOL. Again, this may most likely be due to unspecific labeling as discussed in section 2.6.1. The maximal DOL for SNAPf-tag of 0.36 could be obtained by labeling with 100 nM SiR-BG for an incubation time of 3 h. Alternatively, also an incubation time of 15 min combined with SiR-BG concentrations above 50 nM yielded similarly high DOLs. However, this labeling regime is more prone to unspecific labeling as will be discussed in the following section (compare Fig 2.18). The standard deviations of the DOL of SNAPf-tag in Huh7.5-pBABE-gSEP cells ranged from 0.004 to 0.08 and averaged 0.03.

As mentioned in the previous chapter, random control samples were generated for both protein tags by rotating the coordinates of Halo-TMR and SNAPf-SiR particles 90° clockwise. These samples were used to control for random colocalization that may artificially increase the observed DOL. Using these rotated, random particles and the corresponding unrotated eGFP particles as input for the DOL analysis, both protein tags showed a random colocalization below 0.1 independent of the labeling condition (Fig 2.15 (b)). The mean colocalization based on random colocalization was 0.07 for the HaloTag and 0.05 for the SNAPf-tag. This shows that random colocalization does not appreciably influence the true DOLs of HaloTag and SNAPf-tag.

As a control for unspecific labeling, Huh7.5-pBABE-LynG cells were exposed to the same labeling conditions as Huh7.5-pBABE-gSEP cells and approximately the same amount of replicates per conditions was imaged. As Huh7.5-pBABE-LynG cells express the LynG construct, which contains a eGFP but neither HaloTag nor SNAPf-tag, particles in the Halo-TMR channel and in the SNAP-SiR channel arise from unspecific labeling. The colocalization that was calculated based on unspecific labeling averaged 0.03 for the HaloTag and 0.07 for the SNAPf-tag. Even for very high substrate concentration combined with long incubation times, that are well out of the optimal labeling regime described in section 2.6.2, this unspecific colocalization does not exceed 0.08 for the HaloTag and 0.14 for the SNAPf-tag (Fig 2.15 (c)). This indicates that the true DOL was mostly uninfluenced by unspecific labeling.

As discussed in the introduction, one of the main benefits of determining the DOL of a fluorescent label is the ability to extract quantitative data such as absolute molecule numbers or concentrations from single-molecule microscopy data. Knowing the DOL of HaloTag and SNAPf-tag allows to determine the absolute

protein concentration of the gSEP construct in two independent ways, namely via the concentration of SNAPf-tag or via the concentration of HaloTag, respectively. For this, I determined the density of Halo-TMR and SNAP-SiR particles in all Huh7.5-pBABE-gSEP cells that were labeled with a substrate concentration between 1–50 nM, which totaled 187 cell images. The restriction on substrate concentration was introduced in order to reduce the influence of insufficient and unspecific labeling as will be discussed in section 2.6.1. For all cells and both tags individually, I divided the measured particle densities by the corresponding DOL to obtain the absolute particle concentrations of the two tags. Since SNAPf-tag and HaloTag are fused in a 1:1 ratio within the gSEP construct, the two independently measured concentrations should be very similar and eventually represent the concentration of the gSEP construct in the basal plasma membrane. Using either SNAP-SiR particles or Halo-TMR particles as a reference, the concentration of the gSEP construct was found to be 3.7 ± 1.1 particles/ μm^2 and 3.4 ± 1.1 particles/ μm^2 , respectively (Fig 2.16 (a)). Although the estimation based on SNAPf-tag particles obtained a slightly higher gSEP concentration, which is even significant with a p -value of $p = 0.0014$ (Wilcoxon ranksum test), the two concentrations were still well comparable. Furthermore, a significant correlation with a Pearson correlation coefficient of $R = 0.79$ ($p < 0.001$) was found between the absolute concentrations of SNAPf-tag and HaloTag on a single-cell level (Fig 2.16 (b)).

As a summary, the DOL of HaloTag was found to reach a stable plateau above 0.3 for most labeling conditions and to peak at around 0.5. This means that at most half of all present HaloTags were fluorescently labeled, or in other words, the cell membrane contained at least twice as much HaloTags as could be detected in the images. The DOL of SNAPf-tag was found to peak below 0.4 and to be more dependent on the exact labeling conditions when compared to HaloTag labeling. While both tags benefited from long incubation times, the turnover of the target proteins will probably limit the effective incubation time in most live-cell applications. However, overnight labeling of both proteins tags showed the best results for the labeling of interferon- α receptor (IFNAR), epidermal growth factor receptor (EGFR) and hepatocyte growth factor receptor (c-Met) as described in chapter 3 and 4. Finally, to exemplify how determination of the DOL enables quantification, the absolute concentration of gSEP constructs in the basal membrane of Huh7.5-pBABE-gSEP cells could be consistently determined in two independent ways, using measured particle densities and the calculated DOL of either SNAPf-tag or HaloTag.

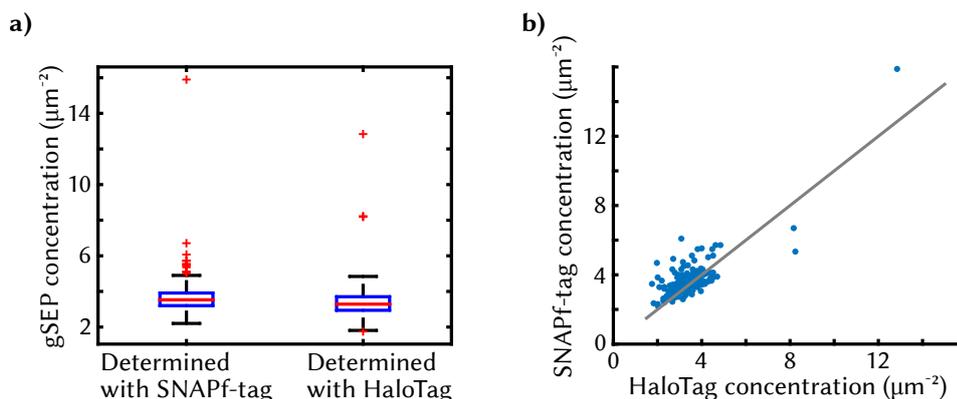


Figure 2.16.: Absolute concentration of the gSEP construct in the basal plasma membrane. The absolute concentration of gSEP constructs was determined for 187 Huh7.5-pBABE-gSEP cell samples labeled with a substrate concentration of 1–50 nM based on SNAPf-tag particles or HaloTag particles. **a)** The two independent estimates based on Halo-TMR particles or SNAP-SiR particles yielded a comparable average gSEP concentration that was slightly higher when estimated via SNAP-SiR particles. **b)** On a single-cell level, the absolute concentration of HaloTag correlated significantly with the absolute concentration of SNAPf-tag (Pearson correlation coefficient $R = 0.79, p < 0.001$). A gray line indicates the expected 1:1 ratio.

2.6.1. Unspecific Labeling of HaloTag TMR ligand and SiR-BG

The high sensitivity of single-molecule microscopy techniques that allows to obtain information from single fluorophores, likewise demands an especially low level of unspecific labeling and background fluorescence. For the determination of the DOL as well as for single-molecule tracking, unspecific labeling can corrupt results by either mimicking target particles or by generating a continuous background intensity that hinders the detection of target particles. Both influences were examined for all tested labeling conditions based on Huh7.5-pBABE-LynG cells. Again, as the LynG construct does neither have the HaloTag nor the SNAPf-tag, it allows to control for unspecifically bound substrate molecules, impurities of the LabTek and autofluorescence of the cell.

To estimate the number of unspecific particles, detected particle densities in Huh7.5-pBABE-LynG cells were determined for all labeling conditions and in all three spectral channels (blue, green and red for eGFP, TMR and SiR detection, respectively). When labeling with HaloTag TMR ligand, an average particle density

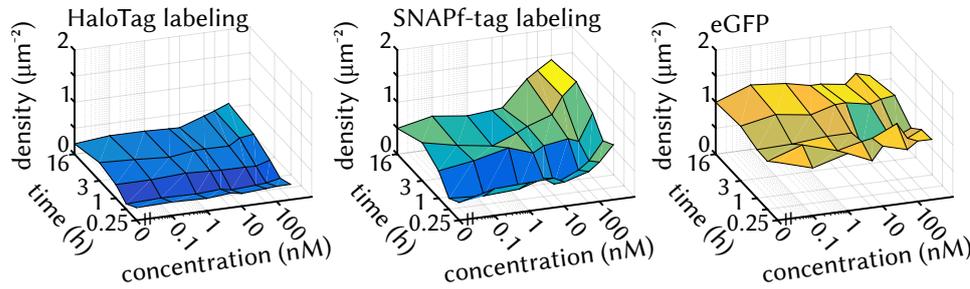
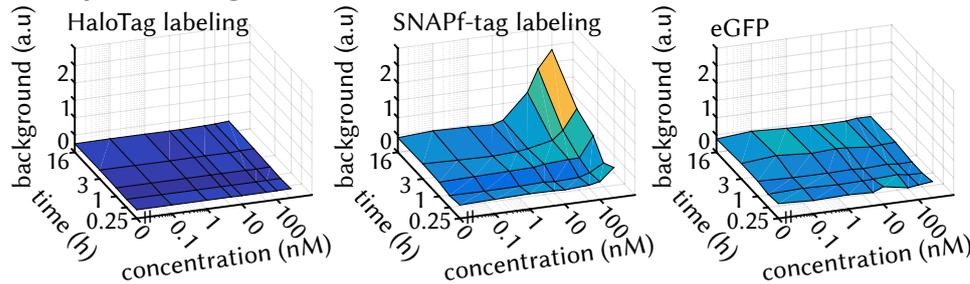
a) unspecific particles**b) unspecific background**

Figure 2.17.: Unspecific labeling of HaloTag and SNAPf-tag. Huh7.5-pBABE-

LynG cells were stained to determine the particle density and the fluorescent background that arises from unspecific labeling **a)** For HaloTag labeling, the density of unspecific particles stayed very low for all labeling conditions except for very high HaloTag TMR ligand concentrations combined with very long incubation times (left). For SNAPf-tag labeling, a higher particle density could be detected partly due to an elevated autofluorescence of Huh7.5-pBABE-LynG cells in the red channel. Still, for SiR-BG concentrations above 10 nM combined with long incubation times, the density of unspecific particles increased strongly and peaked at values well below the detected density of eGFP particles (center, right). **b)** HaloTag labeling produced only minimal fluorescence background for all labeling conditions (left). In contrast, SiR-BG created a strong unspecific background at concentrations above 50 nM and incubation times longer than 1 h (center). As expected, unspecific background in the eGFP channel was independent of the labeling conditions (right).

of 0.28 particles/ μm^2 was found in the green spectral channel. This is comparable to the average particle density of 0.23 particles/ μm^2 found in unstained cells, indicating that most of these particles arise from impurities or residual dirt instead of unspecifically bound HaloTag TMR ligand molecules. Only for HaloTag TMR

ligand concentrations above 50 nM combined with incubation times of 3 h or more, the density of unspecific particles increases markedly to 0.4–0.6 particles/ μm^2 (Fig 2.17 (a), left). However this values is still well below the average density of eGFP particles of 1.07 particles/ μm^2 (Fig 2.17 (a), right).

For SNAPf-tag labeling, detected particle densities were influenced by a strong autofluorescence of Huh7.5-pBABE-LynG cells in the red channel. Even unlabeled cells showed a comparably high particle density of 0.46 particles/ μm^2 . However, for SiR-BG concentrations above 10 nM combined with incubation times of 1 h or longer, the density of unspecific particles increased far above this level to up to 1.42 particles/ μm^2 . Notably, SiR-BG concentrations above 50 nM created an increased density of unspecific particles independent of the incubation time. For these concentrations the tendency of SiR-BG for unspecific labeling seemed to be so high that it can hardly be counteracted by shortening the incubation time. Such a threshold behavior could not be detected for the incubation time. Overnight labeling did not automatically generate a high unspecific labeling but only lead to a high particle density for concentrations of 50 nM or higher.

In addition to the number of unspecific particles, also the level of continuous background fluorescence was estimated. The fluorescent background of Huh7.5-pBABE-LynG cells was determined as the average background intensity of the segmented cell area. The required background intensity was estimated using a rolling ball algorithm with a radius of 20 pixel. For unstained cells, an average unspecific background of 0.41, 0.25 and 0.40 was found for the blue (eGFP), green (TMR) and red (SiR) channel, respectively (Fig 2.17 (b)). The differences between these values can be explained by different absolute laserpowers and different amounts of autofluorescence in the three spectral channels. For the blue channel, also the presence of specific eGFP signals by tendency increases background intensity detected by the rolling ball algorithm. As could be expected, the amount of fluorescent background in the eGFP channel was constant for all labeling conditions. Notably, also the amount of fluorescent background in the TMR channel was largely independent of the staining condition. Only for overnight labeling with 250 nM HaloTag TMR ligand, the background increased slightly to 0.30. In contrast, the unspecific background fluorescence in the SiR channel increased strongly for SiR-BG concentrations above 50 nM. This effect was further enhanced by increasing the incubation times above 1 h.

For both protein tags, the increase of unspecific labeling can be a possible explanation why the DOL showed a tendency to decrease for high substrate concentrations (Fig 2.15 (a)). As the particle detection is based on a two-dimension Gaussian intensity fit, it is restricted by high particle densities as well as a high background fluorescence. While the DOL was corrected for the detected particle density (see chapter 2.4 on page 35 for details), a high unspecific background

fluorescence can further limit the DOL by masking specific particles.

As a summary, these results indicate that SiR-BG is far more likely to produce unspecific labeling than HaloTag TMR ligand. Based on these findings, I would discourage the use of SiR-BG concentrations above 50 nM for single-molecule experiments as a considerable amount of unspecific particles and unspecific background fluorescence can be expected. For incubation times above 1 h, the SiR-BG concentration should be even further reduced. HaloTag TMR ligand on the other hand, allowed for a wide range of concentrations and incubation times without producing considerable levels of unspecific labeling. The increase in unspecific labeling for high substrate concentrations and long incubation times is considerably smaller and limited to more extreme labeling conditions for HaloTag TMR ligand compared to SiR-BG.

2.6.2. Optimal Labeling of HaloTag and SNAPf-tag for Single-Molecule Microscopy

In the preceding sections, I have determined the DOL as well as the amount of unspecific labeling under varying labeling conditions for SNAPf-tag and HaloTag. Very generally, both the DOL as well as the amount of unspecific labeling increases with concentration and incubation time. Hence, a tradeoff has to be made between efficient labeling on the one side and unspecific labeling on the other side. A good labeling condition should combine a high DOL with a low amount of unspecific labeling.

For HaloTag labeling, these two requirements were met by a relatively wide range of labeling conditions (Fig 2.18 right column, range of recommended labeling conditions highlighted in blue). Overnight labeling with HaloTag TMR ligand concentrations between 0.1–10 nM gave the best tradeoff and should be used if incubation time is not limiting. For shorter incubation times, HaloTag TMR ligand concentrations of 1–50 nM resulted in an only slightly lower DOL, while hardly any unspecific labeling was observed.

For SNAPf-tag labeling, the sweet spot for labeling was found to be considerably smaller (Fig 2.18, left column). For short incubation times, a high DOL could only be achieved with SiR-BG concentrations that created a considerable level of unspecific staining at the same time. However, incubation times of at least 3 h allowed to use SiR-BG concentrations in a range of 1–10 nM to achieve a high DOL while levels of unspecific labeling remained low. A combination of 50–250 nM SiR-BG and a incubation time of 15 min, which was identified in section 2.6 to yield a high DOL, is utilizing a concentration range where an increase in unspecific labeling is detectable even for such short incubation times. For single-molecule microscopy, this regime is therefore not recommended.

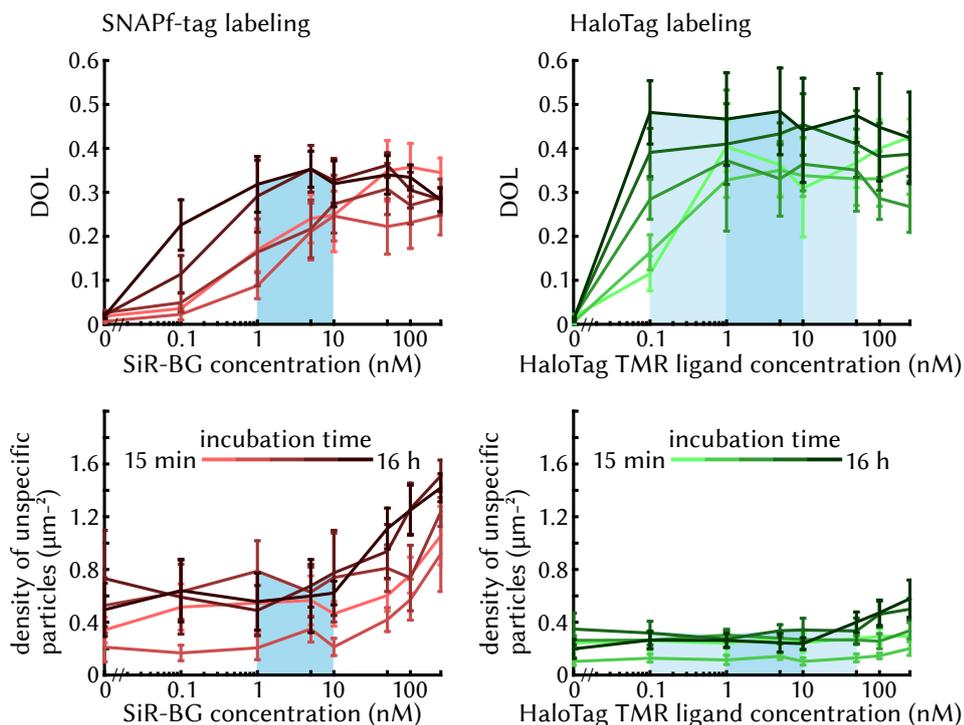


Figure 2.18.: optimal labeling conditions for SNAPf-tag and HaloTag. For single-molecule microscopy based on SNAPf-tag and HaloTag fusions, an optimal labeling protocol should result in a large DOL while keeping unspecific labeling as low as possible. A good tradeoff for SNAPf-tag labeling combined low SiR-BG concentrations of 1–10 nM with incubation times above 3 h (left column, highlighted in blue). Efficient HaloTag labeling with low unspecific labeling can be achieved with HaloTag TMR ligand concentrations of 1–10 nM (right column, highlighted in blue). Slightly balancing HaloTag TMR ligand concentration and incubation time easily increases this range to 0.1–50 nM HaloTag TMR ligand (right column, highlighted in light blue). All errors are standard deviations from at least 10 replicates

To sum up, overnight labeling with 1–10 nM dye substrate was found to yield a high DOL and low unspecific labeling for both protein tags. Generally, reducing the substrate concentration effectively limited unspecific labeling while long incubation times ensured a high degree of labeling at the same time.

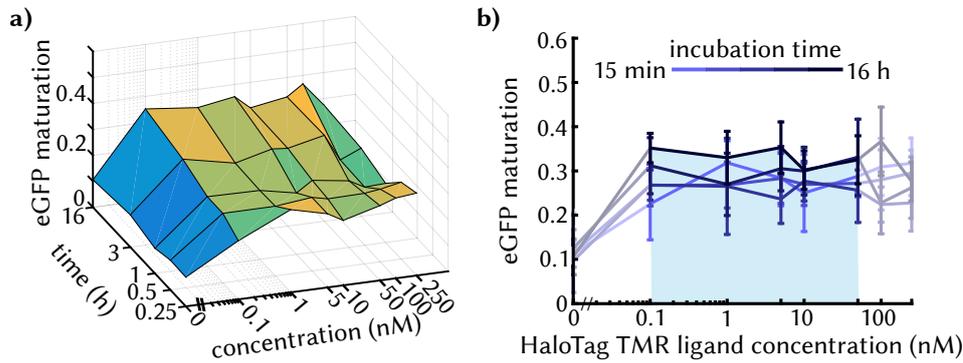


Figure 2.19.: Degree of eGFP maturation. The degree of eGFP maturation was determined based on a colocalization analysis of eGFP particles and Halo-TMR particles in stained Huh7.5-pBABE-gSEP cells. For efficient and largely background-free HaloTag labeling conditions (highlighted in light blue), the degree of eGFP maturation was relatively robust and averaged 0.29 ± 0.06 . Errors are standard deviations from at least 10 replicates.

2.6.3. Degree of Maturation of eGFP

The main reason to incorporate eGFP into the gSEP construct was to have a reference signal that is practically background free. This however was found to be also true for most HaloTag labeling conditions. In principle, it is therefore possible to determine the “DOL” of eGFP based on its colocalization with Halo-TMR. The DOL of eGFP can be better described as the degree of maturation, which is the fraction of all eGFPs that actually fold into a functional fluorescent form. With regard to equation 2.1, this maturation rate can be calculated as

$$DOL_{eGFP} = \frac{|eGFP \cap Halo-TMR|}{|Halo-TMR|} \quad (2.7)$$

,where $eGFP$ is the set of eGFP particles and $Halo-TMR$ is the set of Halo-TMR particles.

As expected, the degree of eGFP maturation is largely independent of the HaloTag labeling conditions (Fig 2.19 (a)), as long as a sufficient labeling of HaloTag is achieved. When averaged over all HaloTag labeling conditions with low unspecific labeling (all labeling conditions with a HaloTag TMR ligand concentration between 0.1–50 nM) the degree of eGFP maturation was found to be 0.29 with a standard deviation of 0.06 (Fig 2.19 (b)). This means that below one third of

eGFP proteins actually folded into a functional, fluorescent state. Interestingly, this indicates that not only for protein tags but also for fluorescent proteins it may be critically required to determine the fraction of functional labels if quantitative information is to be extracted from microscopy data.

Chapter 3.

Two-Color Single-Molecule Tracking of Type I Interferon Receptor

The type I Interferon receptor is a well described cytokine receptor that plays a major role in antiviral cell response (see section 1.3.1 on page 11 of the introduction). Because much is already known about the diffusion and interaction dynamics of Interferon- α receptor 1 (IFNAR1) and Interferon- α receptor 2 (IFNAR2) in live cells, the IFNAR system is a valuable validation system to set up a two-color single-molecule tracking routine. In contrast to the most recent publication by Wilmes et al. [30], who quantified the interaction of both receptors at room temperature using relatively heavy quantum dots as fluorescent markers, I performed my measurements at 37 °C and used light protein labels (HaloTag and SNAPf-tag) for more physiological relevance (Fig 3.1 (a)).

To benchmark the dynamic range of dimerization detection, I used the GFP-enhanced Staining Efficiency Probe (gSEP) introduced in section 2 on page 2 as a dimer and monomer control. When labeled with spectrally different SNAPf-tag and HaloTag dye substrates, this artificial protein resembles a stable dimer (Fig 3.1 (b)), while staining only one of the tags with two competing dye substrates results in strict monomers (Fig 3.1 (c)). As the gSEP construct is anchored to the inner leaflet of the plasma membrane due to its myristoylation and palmitoylation sequence without any further complexation or attachment to the cytoskeleton, it is considerably faster than membrane-bound cell receptors and was therefore imaged at room temperature in order to slow down its diffusion. As a major quality check, I could detect strong dimerization in the dimer control and also confirm ligand-induced heterodimerization of IFNAR1 and IFNAR2, however with largely reduced dimer stability than reported by Wilmes et al..

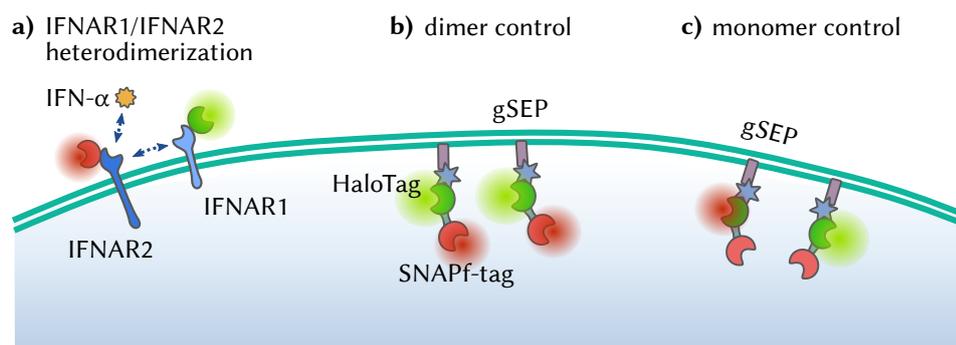


Figure 3.1.: Validation systems for the two-color single-molecule tracking routine. **a)** To test ligand-induced dimerization, IFNAR1 and IFNAR2 were fused to HaloTag and SNAPf-tag, respectively and tracked in presence or absence of IFN- α . **b)** Containing a fusion between SNAPf-tag and HaloTag, the gSEP construct can be used as monomer or dimer control depending on the labeling. When the gSEP construct is labeled with two spectrally distinct dye substrates, it resembles a stable dimer. **c)** Labeling HaloTag with two competing dye substrates generates strict monomers.

3.1. Generating a Stable Huh7.5-Halo-IFNAR1/SNAPf-IFNAR2 Cell Line

In order to simultaneously track both receptor chains in live cells, I stably introduced n-terminal fusion constructs of IFNAR1 and IFNAR2 with HaloTag and SNAPf-tag, respectively, into the hepatocarcinoma cell line Huh7.5. As mentioned in section 1.2.3 on page 7 and demonstrated in chapter 2, these protein tags allow for an orthogonal labeling in live cells using organic dye substrates.

First I transferred Halo-IFNAR1 and SNAPf-IFNAR2 from a pSEMS vector into a retroviral pMOWS vector using XhoI and EcoRI restriction sites in pSEMS and NdeI and EcoRI restriction sites in pMOWS in a one-sided blunt ligation reaction. pSEMS-Halo-IFNAR1 and pSEMS-SNAPf-IFNAR2 were kindly provided by Ursula Klingmüller¹ and originally obtained from Jacob Piehler². The resulting pMOWS-Halo-IFNAR1 and pMOWS-SNAPf-IFNAR2 plasmids were validated by sequencing. Both plasmids were then introduced into Huh7.5 cells via retroviral transduction using the Phoenix Ampho system as described in subsection 6.2.1 on page 113 of material and methods. Co-transduced Huh7.5-Halo-IFNAR1/SNAPf-

¹German Cancer Research Center, Heidelberg

²University of Osnabrück

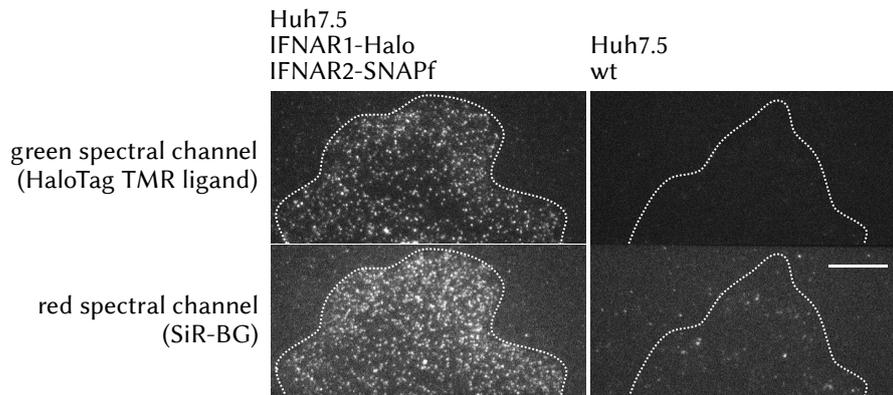


Figure 3.2.: Fluorescent labeling of Halo-IFNAR1 and SNAPf-IFNAR2. TIRFM images of labeled Huh7.5 cells. Live wt Huh7.5 cells and Huh7.5-Halo-IFNAR1/SNAPf-IFNAR2 cells were stained with 10 nM SiR-BG and 10 nM HaloTag TMR ligand overnight and imaged under TIRF conditions. Halo-IFNAR1 and SNAPf-IFNAR2 could both be specifically labeled in Huh7.5-Halo-IFNAR1/SNAPf-IFNAR2 cells (left), while hardly any signal was detectable in wt cells (right). Scale bar is 10 μm . For comparability, the intensities are consistent for all images.

IFNAR2 cells were selected with $1.5 \mu\text{g mL}^{-1}$ puromycin and $400 \mu\text{g mL}^{-1}$ G418 over two weeks.

To check the expression and correct localization of both constructs, I recorded live-cell fluorescence images of both constructs under total internal fluorescence reflection (TIRF) conditions. Building on the gained knowledge about the optimal labeling conditions of SNAPf-tag and HaloTag presented in section 2.6.2, I used an overnight labeling procedure for live Huh7.5-Halo-IFNAR1/SNAPf-IFNAR2 cells based on the orthogonal dye substrates Halo-tetramethylrhodamine ligand (HaloTag TMR ligand) and siliconrhodamine-benzylguanine (SiR-BG). More specifically, I seeded wt Huh7.5 cells and Huh7.5-Halo-IFNAR1/SNAPf-IFNAR2 cells in a cleaned Labtek and stained each cell line overnight with 10 nM SiR-BG and 10 nM HaloTag TMR ligand as described in section 6.8.1 on page 121 of material and methods.

This labeling procedure resulted in a specific and orthogonal labeling of Halo-IFNAR1 and SNAPf-IFNAR2 while producing hardly any unspecific labeling in wt cells (Fig 3.2). Membrane-bound Halo-IFNAR1 and SNAPf-IFNAR2 could be clearly detected in Huh7.5-Halo-IFNAR1/SNAPf-IFNAR2 cells. This confirmed a specific labeling of both constructs as well as their correct expression and localization.

3.2. Two-Color Single-Molecule Tracking of IFNAR1 and IFNAR2

To determine diffusion and interaction properties of the two receptor chains, I developed and implemented the two-color single-molecule tracking work-flow depicted in Figure 3.3. After specifically labeling both receptor chains, I imaged them on a TIRF microscope at 37 °C using a custom-build heating chamber and an objective heating ring. To be able to simultaneously image both receptors, I projected each spectral channel to half of the camera chip. Before each tracking experiment, I imaged multicolor beads as fiducial markers visible in both spectral channels and used them to derive a transformation matrix for image registration.

For the receptors, I performed single-molecule tracking in each spectral channel and projected the obtained trajectories into a common coordinate space using the previously obtained transformation matrix. From all trajectories, I selected trajectory pairs that showed a temporal and spatial overlap, as this is a prerequisite for interaction. For each pair of candidate trajectories, I calculated the distance between the corresponding particles in each frame. Using a fixed distance threshold, I classified each timestep of the trajectory pair as being either a “free” or a “dimer” state. From the durations of dimer states, I obtained the dissociation rate of the IFNAR complex. In addition, I analyzed the jump sizes of free and dimer receptors as well as the total fraction of heterodimers.

Using this routine, I could well discriminate the positive and negative dimer controls and also verify heterodimerization of IFNAR1 and IFNAR2 upon stimulation with IFN- α 2. As ligand-induced heterodimerization between IFNAR1 and IFNAR2 is well documented in literature and reported on a single molecule level [30, 80], detecting IFNAR1/IFNAR2 heterodimers was a crucial quality control for the two-color single-molecule tracking setup. In addition, I could observe reduced dimerization at very high IFN- α 2 concentrations, a finding that has been supported by a kinetic model developed by Nikolas Schnellbacher³.

In the following I will describe the individual steps of the two-color single-molecule tracking routine and the obtained results in more detail.

3.2.1. Single-Molecule Tracking and Registration of Spectral Channels

As depicted in Figure 3.3, Halo-IFNAR1 and SNAPf-IFNAR2 were imaged in separate spectral channels that were each allocated to half of the camera chip. Single-molecule tracking was then performed in each channel separately. For both re-

³AG Schwarz, Heidelberg University

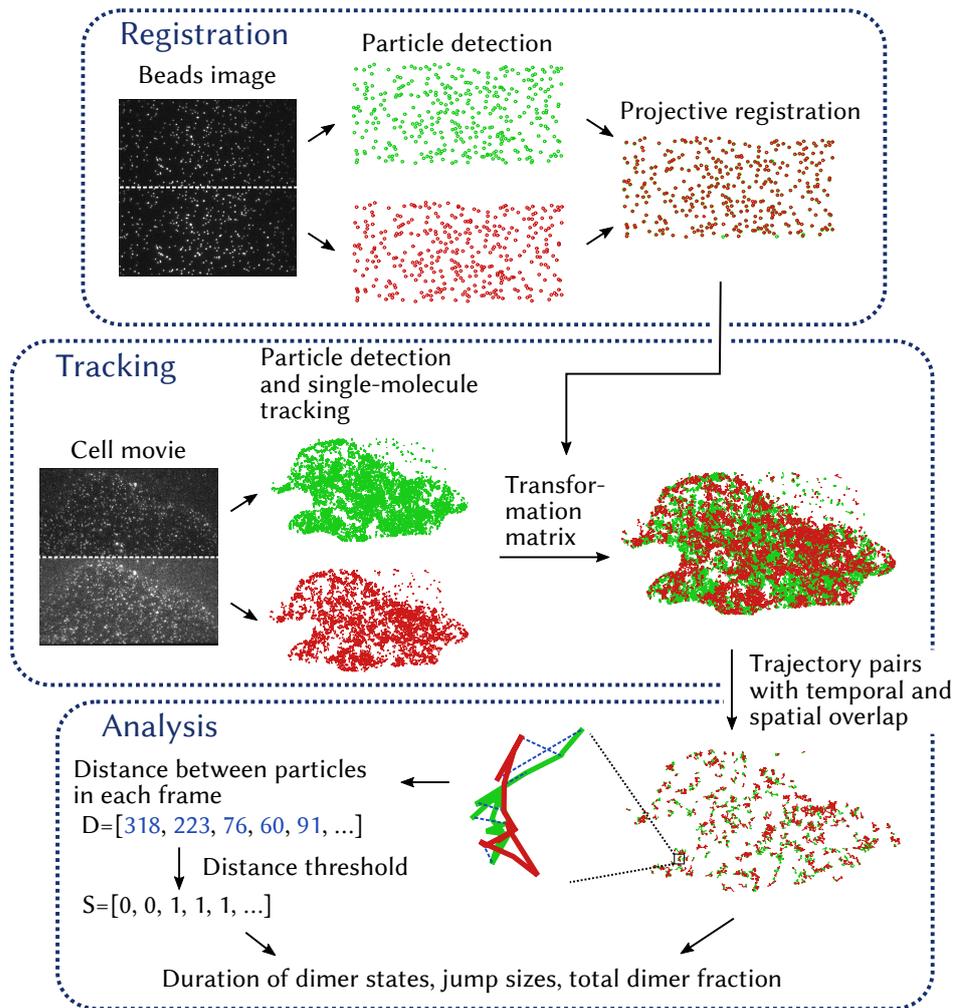


Figure 3.3.: Overview of the two-color single-molecule tracking work-flow

Labeled Halo-IFNAR1 and SNAPf-IFNAR2 were simultaneously imaged on a split camera chip. Single-molecule tracking was performed for each spectral channel (green for Halo-TMR and red for SNAPf-SiR). The resulting trajectories were registered using a transformation matrix that was obtained from fiducial bead markers. Trajectory pairs with a temporal and spectral overlap were extracted for further analysis. For each trajectory pair and timestep, particles were classified as being in a “free” or “dimer” state using a fixed distance threshold. Based on the measured duration of dimer states, the dissociation rate was determined. Additional analysis parameters included the observed jump sizes and the total fraction of heterodimerization.

ceptors, I used the Point Source Detection algorithm described in section 2.4 to detect single molecules, however with allowing to fit mixtures of Gaussian intensities profiles. Using mixed Gaussian fits increases the sensitivity of the particle detection at the cost of a higher fraction of false positive detection as discussed in section 2.4. Because false positive detections appear more stochastically than true positive detections, they hardly give rise to trajectories. Hence, in the context of single-molecule tracking, this modification yields a positive trade-off, while for pure single-molecule detection in fixed cells it proved disadvantageous (see section 2.4). In order to limit detection to the area of the cell, a segmentation based on the SNAPf-IFNAR2 signal was performed as described in section 2.3.1 on page 32. To utilize the mobility of SNAPf-IFNAR2 particles for segmentation, a max projection of all frames was used as input for the segmentation. In this way, bright particles leave an intensity trace as they sample the cell over time, which facilitates segmentation of cells with a low particle density. Formation of trajectories from detected particles was performed using u-track with default parametrization.

In order to align the particle trajectories of Halo-IFNAR1 and SNAPf-IFNAR2, I developed an image registration based on fiducial markers that were recorded before the tracking experiment. In a first step, a 512x250 pixel area centered at each half of the camera chip was extracted (Fig 3.4 (a), left) and used as input for the point detection step. From all detected points, I extracted those that had exactly one partner in the other channels within a search radius of 10 pixel when both channels were simply overlaid (Fig 3.4 (a), center). These non-ambiguous particles pairs were used as landmarks to calculate a projective transformation matrix using the MATLAB function “fitgeotrans”.

To evaluate the success of the registration, I calculated the Euclidean nearest neighbour distance between particles of the two spectral channels before and after registration for a total of 23 individual fiducial marker images (Fig 3.4 (b)). After projective registration, the median nearest neighbour distance was 47 nm (equivalent of 0.4 pixel) and 95 % of distances were below 136 nm (equivalent of 1.3 pixel). This compared to a median nearest neighbour distance of 322 nm (equivalent of 3.1 pixel) when channels were simply overlaid.

As the fiducial marker-based transformation matrix was used to register IFNAR1 and IFNAR2 trajectories, its stability had to be determined over the duration of an experiment. This was especially important since considerable variations of registration parameters between samples were observed during the determination of the degree of labeling of HaloTag and SNAPf-tag earlier in this study (see section 2.5.2 on page 43). For a subset of tracking experiments, I therefore recorded fiducial marker images before and after the experiment. These two timepoints differed by approximately 2 h. I then calculated a transformation matrix based on an early recorded image and used it to either register the corresponding image (Fig 3.4 (c),

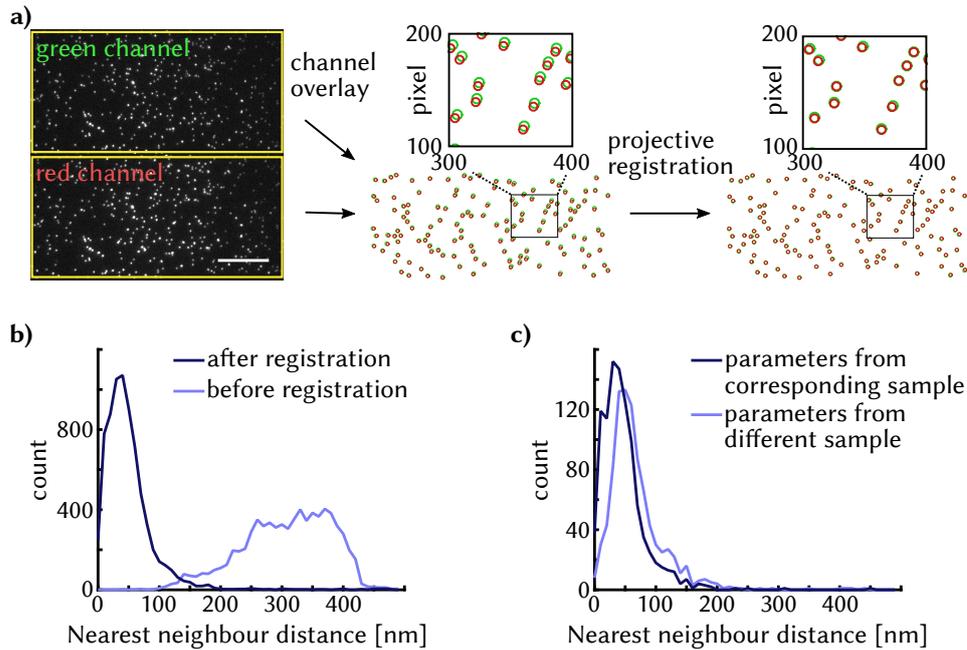


Figure 3.4.: Image registration to align the two spectral channels based on fiducial markers. **a)** In TIRM images of multicolor beads, particles were detected in 512×250 pixel areas representing the center of each spectral channel. Non-ambiguous particle pairs were used to determine a projective transformation matrix. **b)** The Euclidean distance between nearest neighbour particles of different channels was calculated for 23 individual beads images with a total of 7436 particle pairs. Projective registration decreased the median nearest neighbour distance from 322 nm to 47 nm. **c)** Fiducial marker images were recorded before and after a tracking experiment. Based on three pairs of images, a median nearest neighbour distance of 47 nm was confirmed when images were registered with their own transformation matrix (1051 distances, dark blue). When using the transformation matrix of an early recorded image to register a late recorded image, the median nearest neighbour distance was slightly increased to 62 nm (924 distances, light blue).

dark blue) or a different, later recorded image (Fig 3.4 (c), light blue). In total, three pairs of images were analyzed.

For images that were registered with their corresponding transformation matrix, the median nearest neighbour distance of 47 nm was confirmed. When the transformation matrix from an early recorded image was used to register a later

recorded image, the median nearest neighbour distance was slightly increased to 62 nm (equivalent of 0.6 pixel) and 95 % of distances were below 160 nm (equivalent of 1.5 pixel). This increase and the resulting nearest neighbour distances were considered small enough to allow for dimerization detection in two-color single-molecule tracking data.

As a result, these experiments showed that a projective image registration effectively aligned fiducial marker images and that the temporal robustness of the so derived transformation matrices allows to use fiducial marker-derived transformation matrices to align receptor trajectories.

3.2.2. Estimation of a Distance Threshold for Dimer Classification

Even after successful registration, most of the detected Halo-IFNAR1 and SNAPf-IFNAR2 molecules have no chance to interact because they are too distant to each other or because they appear during different times of the movie. From all registered Halo-IFNAR1 and SNAPf-IFNAR2 trajectories, I therefore extracted “candidate trajectory pairs” that fulfilled two requirements: They had to show a temporal overlap of at least 10 frames (which equals 0.2 s) and they had to be closer than 1.8 pixel (equivalent to 187 nm) for at least one frame.

To exclude artificial dimer-trajectories that may result from the detection of cellular autofluorescence or dirt particles, an additional mobility threshold was introduced that used the mean jump size of a trajectory. The jump size describes the mean Euclidean distance of a particle in two consecutive frames of a trajectory. As cellular autofluorescence and dirt particle are relatively immobile compared to diffusing receptors, a mobility threshold was estimated by tracking immobile multicolor beads. For the resulting trajectories, a mean jump size distribution was calculated and a threshold of 0.7 pixel (equals 74 nm) was determined in order to reject 95 % of beads trajectories (Fig 3.5 (a)). Only candidate trajectories with a mean jump size above this threshold were further analyzed for dimerization events.

A widely used classification criterion for particle interaction is the distance between them. In an optimal system, the distance between the two dimerization partners represents their actual molecular distance given by the molecular structure of the complex and the position and size of the labels. In diffraction limited multi-color systems however, the observed distance between dimerization partner is mainly determined by the localization and registration error. As shown in the previous section, these errors introduced a median distance between fiducial markers in the range of 60 nm for the here developed system. Because single-particle tracking in densely labeled cells and the optical properties of the cells themselves may by tendency further increase the localization and registration

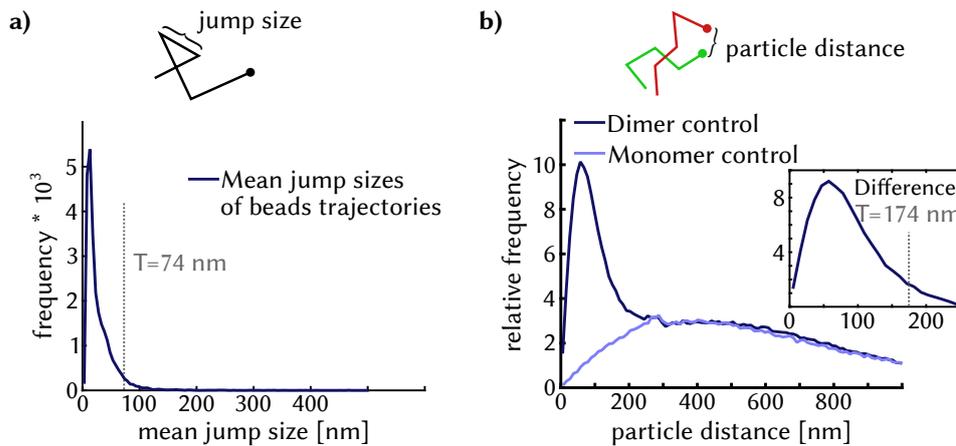


Figure 3.5.: Estimation of a velocity and a distance threshold. **a)** Immobile multicolor beads were tracked and the mean jump size was calculated for each trajectory ($n = 25970$). 95% of trajectories showed a mean jump size below 74 nm, i.e. 0.7 pixel (dotted line). Hence, this velocity threshold was used to discard immobile particle trajectories. **b)** Histogram of distances between candidate particles in the dimer control (dark blue, $n = 479873$) and the monomer control (light blue, $n = 173085$). Both histograms were normalized to their respective count at a distance of 1 μm . In the dimer control, particle distances around 74 nm are highly abundant, while the monomer control showed more equally distributed distances and few distances below 200 nm. The difference between the two distributions yielded the specific dimer distances. In order to correctly classify 95% of dimers based on their distance, a threshold of 174 nm (equivalent to 1.7 pixel) needed to be applied (dotted line).

error, the distance between receptor dimers could be expected to be in the range of or slightly larger than 60 nm.

To quantify the expected distance between dimer particles under the conditions of live-cell single-molecule tracking, I determined the distance between the HaloTag and SNAPf-tag signal of corresponding gSEP constructs in Huh7.5-pBABE-gSEP cells. By labeling HaloTag and SNAPf-tag with different dye substrates, I created a dimer control, while labeling only HaloTag with two competitive dye substrates represented a monomer control (Fig 3.1 (b) and (c) on page 70). For the dimer control, I seeded Huh7.5-pBABE-gSEP cells in a cleaned Labtek and stained them overnight with 10 nM HaloTag TMR ligand and 10 nM SiR-BG. For the monomer control, I stained Huh7.5-pBABE-gSEP cells overnight with

10 nM HaloTag TMR ligand and 50 nM SiR-Halo. For both controls, I performed live-cell single-molecule tracking, registered the obtained trajectories using fiducial markers and extracted candidate trajectory pairs as described above. From these candidate pairs I selected those that showed a distance below a very tolerant distance threshold of 1 μm for more than 5 frames and determined the distance between the partner particles for all frames where this distance threshold was met. For comparability between the controls, I normalized the resulting distance distributions to the frequency of an observed distance of 1 μm as such a large particle distance can be assumed to be independent of dimerization events.

As shown in Figure 3.5 (b), the observed distances between candidate particles in the monomer control are widely distributed and only few particle pairs showed a distance below 200 nm. For the dimer control in contrast, there was a high abundance of particle distances below 200 nm, while for larger distances the distribution approached the monomer control. I assumed that this peak related to double-stained gSEP constructs that showed colocomotion in both spectral channels. To specifically separate these 'stable dimers' from randomly interacting particles I subtracted the monomer distance distribution from the dimer distance distribution (Fig 3.5 (b), 'Difference'). The resulting distance distribution of co-diffusing gSEP constructs showed a median distance of 74 nm (equivalent to 0.7 pixel), which is, as expected, in the same range but slightly higher than the distance between fiducial markers as determined in the previous section. 95 % of distances between dimer particles were below 174 nm (equivalent to 1.7 pixel). Hence, to correctly classify 95 % of dimers, a distance threshold of 1.7 pixel was found to be required (Fig 3.5 (b), dotted line). This distance threshold was set for the further analysis of IFNAR1/IFNAR2 heterodimers.

To sum up, I could derive a distance threshold to classify dimerization events under live-cell single-molecule tracking conditions using a monomer and a dimer control. As the determination of this threshold included the complete two-color single-molecule tracking workflow it also integrates the errors of each step and should be well suitable to classify receptor dimerization.

3.2.3. Quantification of IFNAR1/IFNAR2 heterodimerization

With an efficient registration and a live-cell derived distance threshold at hand, I proceeded to quantify dimerization events between Halo-IFNAR1 and SNAPf-IFNAR2 under different stimulation conditions. To induce heterodimerization, Huh7.5-Halo-IFNAR1/SNAPf-IFNAR2 cells were stimulated with either 1 nM or 100 nM IFN- α 2. As controls, I included the above mentioned dimer and monomer controls as well as unstimulated Huh7.5-Halo-IFNAR1/SNAPf-IFNAR2 cells. For all conditions, Halo-IFNAR1 and SNAPf-IFNAR2 trajectories (or gSEP trajectories

for the dimer and monomer control) were registered and candidate trajectory pairs were extracted as described above. For each timestep, the distance between candidate particles was determined and the particles were classified as 'dimer' if they showed a distance below a distance threshold of 1.7 pixel (equivalent to 174 nm). Otherwise, the particles were classified as 'free'. An exemplified trajectory pair that shows several dimer states (marked as blue connections between the two partner trajectories) is depicted in Figure 3.6 (b).

Figure 3.6 (a) highlights candidate trajectories with at least 5 timesteps classified as dimer for exemplary cells of different conditions. As a basic validation of the two-color single-molecule tracking system, these 'stable' dimer trajectories were found to be far more abundant in the dimer control than in the monomer control, where hardly any stable dimer trajectories could be found. In Huh7.5-Halo-IFNAR1/SNAPf-IFNAR2 cells, the number of stable dimer trajectories was by tendency found to increase upon stimulation with 1 nM IFN- α 2.

To quantify the degree of dimerization, I determined the fraction of dimer states for each candidate trajectory pair, e.g. for a candidate trajectory that has 10 out of 30 frames classified as dimer, the dimer fraction would be 33 %. As shown in Figure 3.6 (c) the monomer and dimer controls created a dynamic detection range of 15–36 %. This range is likely limited by random colocalizations and the detection of 'artificial' dimers like vesicles on the lower end, and by the detection, tracking and registration errors on the upper end. Within this range however, the dimerization between Halo-IFNAR1 and SNAPf-IFNAR2 significantly increased from 23 % to 27 % upon stimulation with 1 nM IFN- α 2. When stimulated with 100 nM IFN- α 2, dimerization increased only slightly, yet significant, to 24 %. When projected to the dynamic range of the system (monomer and dimer control set as 0 % and 100 %, respectively), unstimulated Huh7.5-Halo-IFNAR1/SNAPf-IFNAR2 cells showed a heterodimerization of 37 %. When stimulated with IFN- α 2, heterodimerization increased to 55 % for 1 nM IFN- α 2 and 40 % for 100 nM IFN- α 2. For all conditions, the median jump size of dimer particles was significantly lower than the median jump size of free particles, indicating that dimerization reduces the mobility of individual receptors (Fig A.2 (a)-(b) on page 133 of the appendix).

The reduced heterodimerization at highly saturated ligand concentrations is very likely a consequence of the stoichiometry of the ternary signaling complex, where two individual receptor chains bind a single ligand. If the ligand concentration is very high, receptor chains are more likely to form a 1:1 receptor:ligand complex. In this case, even the low-affinity receptor chain IFNAR1 has a high probability to bind its own ligand instead of being recruited into a ternary complex. This behaviour was predicted by a kinetic model developed by Nikolas Schnellbacher⁴

⁴AG Schwarz, Heidelberg University

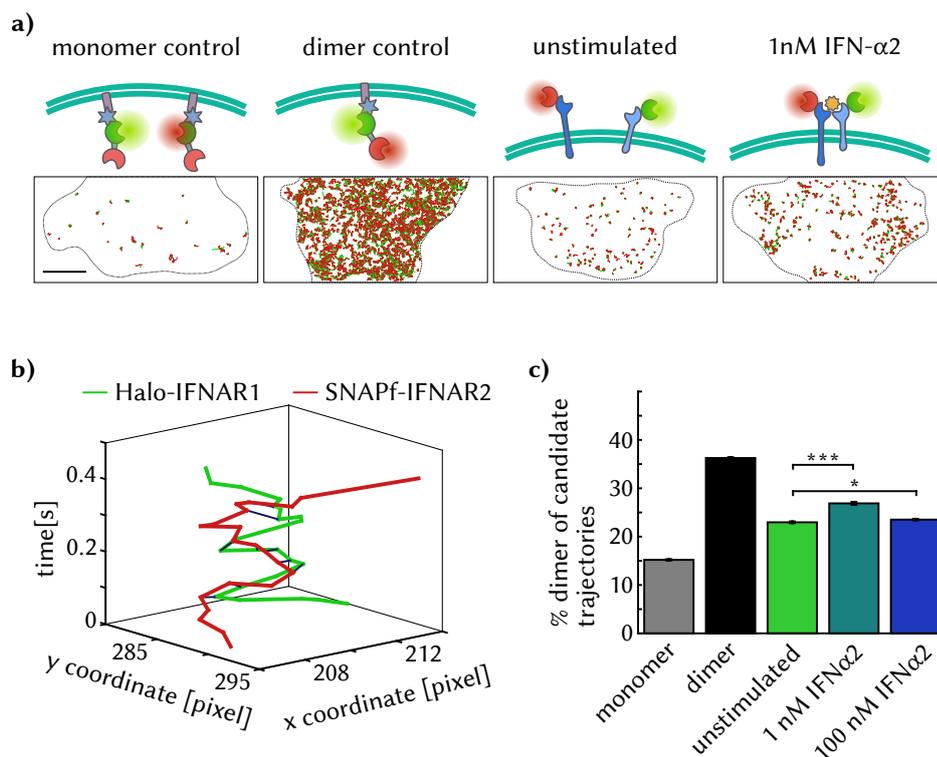


Figure 3.6.: Detection of IFNAR1/IFNAR2 heterodimers. **a)** Visualization of dimer trajectories that show at least 5 frames of 'dimer' state. As expected, the amount of dimer trajectories differed greatly between the monomer and the dimer control. In Huh7.5-Halo-IFNAR1/SNAPf-IFNAR2 cells, stimulation with 1 nM IFN- α 2 by tendency increased the number of dimer trajectories. Scale bar is 10 μ m. **b)** Exemplified trajectory of an IFNAR1/IFNAR2 heterodimer. Dimer states are indicated via blue connections. **c)** Temporal fraction of candidate trajectories classified as 'dimer'. The monomer and dimer control benchmarked a dynamic range between 15–35% dimerization. Within this range, stimulation with 1 nM IFN- α 2 led to a significant increase in IFNAR1/IFNAR2 heterodimerization. In contrast, this effect was considerably weaker upon stimulation with 100 nM IFN- α 2. Errors are standard error of the mean (SEM) and significance levels are $p < 0.05$ (*) and $p < 0.001$ (***) based on bootstrap significance testing (see section 6.8.3 on page 123 of material and methods). The number of analyzed candidate trajectories were 7466 for the monomer control, 23 925 for the dimer control, 11 497 for unstimulated, 5482 for 1 nM IFN- α 2 and 16 063 for 100 nM IFN- α 2.

as shortly discussed in section A.3 on page 134 of the appendix.

To characterize the stability of IFNAR1/IFNAR2 heterodimers, I estimated the dissociation rate from individual dimer durations. For this, I determined the duration of all detected dimer states of individual conditions and estimated the dissociation rate as the exponential decay of the dimer duration histogram (Fig 3.7 (a)). Important to note, this dimer histogram includes two populations, namely 'true' dimerization events as well as random encounters of non-interacting molecules. To disentangle these two, I used a two component exponential fit with one component assigned to random colocalizations while the other component described dimerization events (eq. 3.1).

$$y = b_0 e^{-k_{d,random} \cdot x} + b_1 e^{-k_{d,dimer} \cdot x} \quad (3.1)$$

With the monomer control resembling mostly random colocalizations, I could estimate this first component by fitting the dimer duration histogram of the monomer control with a one component exponential model ($k_{d,random} = 56.9 \text{ s}^{-1}$). I then set this dissociation rate for random interaction as a fixed parameter of a two component exponential model to estimate the specific dimer dissociation rate $k_{d,dimer}$ as second component. Fitting the monomer control with a two component exponential model yielded a second component with low amplitude and large error, indicating overfitting.

Using this approach I could detect a significant stabilization of IFNAR1/IFNAR2 heterodimers upon stimulation with 1 nM IFN- α 2. The dimer specific dissociation rate $k_{d,dimer}$ decreased from 25.4 s^{-1} in unstimulated cells to 22.7 s^{-1} when stimulated. This corresponds to an average lifetime of 44 ms. When stimulated with 100 nM IFN- α 2, only a minor, non-significant decrease of the dissociation constant to 25.0 s^{-1} was observed. This supports the assumption, that under ligand-oversaturation each reception can bind its own ligand. Hence, the ligand loses its cross-linking function and as a result, fully ligand-bound receptor chains show an affinity toward each other which is comparable to the unstimulated situation. The measured IFNAR1/IFNAR2 dissociation rates were well within the dynamic range given by the monomer and dimer control. Hence, it can be assumed that they were not directly limited by the system.

Obviously, the distance threshold for the classification of dimers is a critical parameter for the quantification of dimer events. To test the influence of the distance threshold on dimerization and dimer stability, I performed a parameter variation where I calculated the dissociation rates as well as the degree of dimerization for all conditions in dependence of the distance threshold (Fig 3.7 (c) and (d)).

As expected, choosing a very low distance threshold resulted in a lower degree of dimerization as well as a higher dissociation rate. This is because a very strict

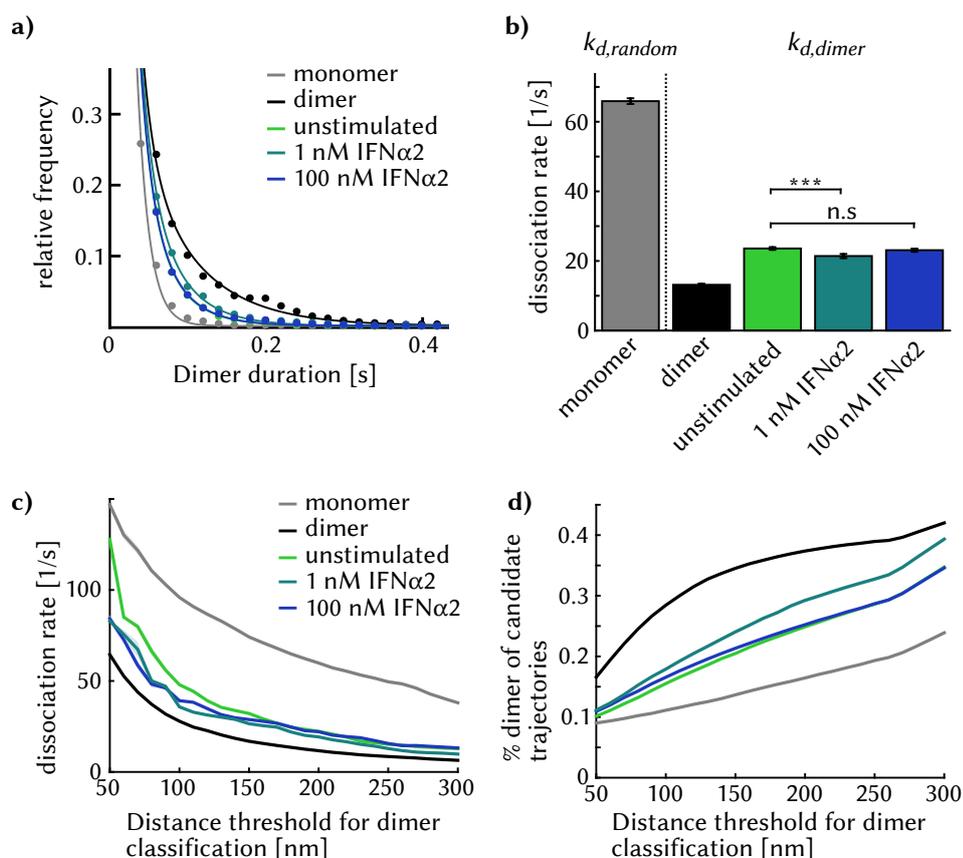


Figure 3.7.: Dissociation rates of IFNAR1/IFNAR2 heterodimers and parameter variation. **a)** Histogram of dimer durations for all conditions. The monomer control was fitted with a one-component exponential model to extract the kinetics of random encounters of non-interacting molecules. For all other conditions, this component was used as a fixed parameter in a two-component exponential fit with the second, free parameter describing the kinetics of dimer dissociation. **b)** First or second component of the exponential fits in a). The dissociation rate of IFNAR1/IFNAR2 heterodimers is significantly reduced upon stimulation with 1 nM IFN- α 2, while stimulation with 100 nM IFN- α 2 led to a considerably smaller reduction of the dissociation rate. Errors are standard errors of the exponential model. Significance level is $p < 0.001$ (***) based on bootstrap significance testing (see section 6.8.4 on page 123 of material and methods). **c)-d)** Parameter variation of the distance threshold for dimer classification. Within a reasonable range between 130–300 nm, the distance threshold influenced the dissociation rate and the degree of dimerization by a factor three and two, respectively. However, the qualitative relations between the individual conditions were conserved and independent of the distance threshold.

dimer classification criterion reduces the number of detected dimer events and therefore also shortens dimer durations. A large dimerization threshold in contrast, led to a high degree of dimerization as well as a low dissociation rate. In respect to the degree of dimerization, a distance threshold above 130 nm was required to reach the full dynamic range of the system as could be assessed by the separation of monomer and dimer control. For smaller distance thresholds, dimer events are likely to be missed.

When increasing the distance threshold in a range from 130–300 nm, the degree of IFNAR1/IFNAR2 heterodimerization differed approximately by a factor of 2 for all stimulation conditions. However, the qualitative differences between the conditions were robust. For all distance thresholds, moderately stimulated cells showed the largest degree of IFNAR1/IFNAR2 dimerization compared to highly stimulated and unstimulated cells. For the latter two, the degree of dimerization was rather comparable.

Similar results were revealed for the dissociation rate. For distance thresholds between 130–300 nm, the dissociation rate was found to vary approximately by a factor of three. In moderately stimulated cells, the dissociation rate of the IFNAR1/IFNAR2 heterodimers was found to vary from 31 s^{-1} to 10 s^{-1} . Again, the qualitative relation between the different conditions stayed conserved. For all distance thresholds above 100 nm, the dissociation rate of unstimulated and highly stimulated IFNAR1/IFNAR2 heterodimers were consistently higher than the dissociation rate of moderately stimulated heterodimers.

As a summary, the two-color single-molecule system proved capable of detecting specific dimerization events both in the dimer control as well as in stimulated Huh7.5-Halo-IFNAR1/SNAPf-IFNAR2 cells. I found that the degree of dimerization as well as dimer stability was increased upon stimulation with moderate concentrations of IFN- α 2, while ligand-oversaturation kept both the degree of dimerization as well as dimer stability to a value close to the unstimulated condition.

Chapter 4.

Ligand-induced Heterodimerization of Epidermal Growth Factor Receptor (EGFR) and Hepatocyte Growth Factor Receptor (c-Met)

As mentioned in section 1.4.2 on page 15, one of the major limitations of tyrosine kinase inhibitors (TKIs) in the treatment of late-stage non-small cell lung cancer (NSCLC) patients is the emergence of resistance in almost all patients. While secondary mutations of epidermal growth factor receptor (EGFR) have been shown to induce TKI resistance in many cases, the amplification of the MET gene and the associated overexpression of hepatocyte growth factor receptor (c-Met) is a less well understood mediator. Based on a dynamic pathway model of EGFR and c-Met signaling fed by time and dose resolved signaling data, the hypothesis of direct EGFR/c-Met heterodimerization as a means of tuning TKI sensitivity emerged in the group of Ursula Klingmüller¹. In collaboration with Florian Salopiata from the lab of Ursula Klingmüller and Yu Qiang from the lab of Karl Rohr², I tested this hypothesis using two-color single-molecule tracking of EGFR and c-MET in the human NSCLC cell line H838. I could show that EGFR and c-Met form heterodimers upon stimulation with EGF and that heterodimerization inhibits the ligand-induced internalization of c-Met, leading to a prolonged signaling. From the mechanistic understanding of EGFR/c-Met interplay, the expression ratio between EGFR and c-Met emerged as a new biomarker to predict TKI responsiveness in NSCLC patients.

¹German Cancer Research Center, Heidelberg

²German Cancer Research Center, Heidelberg and Heidelberg University

4.1. Generating a stable H838-EGFR-SNAPf/c-Met-Halo cell line

In order to orthogonally label and track epidermal growth factor receptor (EGFR) and hepatocyte growth factor receptor (c-Met), Florian Salopiata fused SNAPf-tag and HaloTag to the c-termini of EGFR and c-Met, respectively and introduced both constructs into H838 cells by retroviral transduction (see section 6.2.1 on page 113 in material and methods).

For both constructs, the retroviral pBABE vector was used as a backbone. The cDNA sequence of human EGFR was obtained via reverse transcriptase of messenger RNA (mRNA) extracted from the NSCLC cell line H1975 and subsequent PCR^{3,4}. The resulting PCR product was subcloned into a pMOWS plasmid using XhoI and PacI restriction sites. EGFR sequence was then placed upstream of SNAPf-tag into a pBABE vector by digesting pMOWS-EGFR with XhoI and AseI and pBABE-gSEP with EcoRI and NdeI. EGFR and the pBABE-SNAPf backbone were then combined by converting EcoRI and XhoI to blunt ends and using the compatible cohesive ends of AseI and NdeI for ligation. In addition, the puromycin resistance was replaced by neomycin resistance using HindIII and AfeI restriction sites.

To generate pBABE-c-Met-Halo, cDNA of HaloTag was first amplified with EcoRI and NdeI overhangs by PCR^{5,6} and then introduced into a c-Met-YFP construct, provided by Roche. The resulting c-Met-Halo was amplified with XhoI and MfeI overhangs by PCR^{7,8} and introduced into a pBABE-puro vector using XhoI and EcoRI restriction sites. Both pBABE-EGFR-SNAPf and pBABE-c-Met-Halo plasmids were verified by sequencing.

To generate stable cells lines, the validated pBABE-EGFR-SNAPf and pBABE-c-Met-Halo plasmids were simultaneously introduced into H838 cells via retroviral transduction using the Phoenix Amphi system as described in subsection 6.2.1. Co-transduced cells were selected with $1.5 \mu\text{g mL}^{-1}$ puromycin and $400 \mu\text{g mL}^{-1}$ G418.

The expression and functionality of both EGFR-SNAPf and Met-Halo constructs was verified by Florian Salopiata via immunoblotting of total and phosphorylated protein. In the cytoplasmic fraction of both wildtype (wt) H838 cells as well as H838-EGFR-SNAPf/c-Met-Halo cells, endogenous EGFR and c-Met could be

³fwd: 5'-CGCCTCGAGATGCGACCCTCCGGGACGG-3'

⁴rev: 5'-GCGTTAATTAATCATGCTCCAATAAATTCAGTCTTGTG-3'

⁵fwd: 5'-GCGGAATTCGCAGAAATCGGTACTGGCTTTCCATT-3'

⁶rev: 5'-CGCCATATGGCCGAAATCTCCAGCGTCGA-3'

⁷fwd: 5'-GCGCTCGAGATGAAGGCTCCTGCCGTGCTG-3'

⁸rev: 5'-CGCCAATTGTTAGCCGAAATCTCCAGCGTCGA-3'

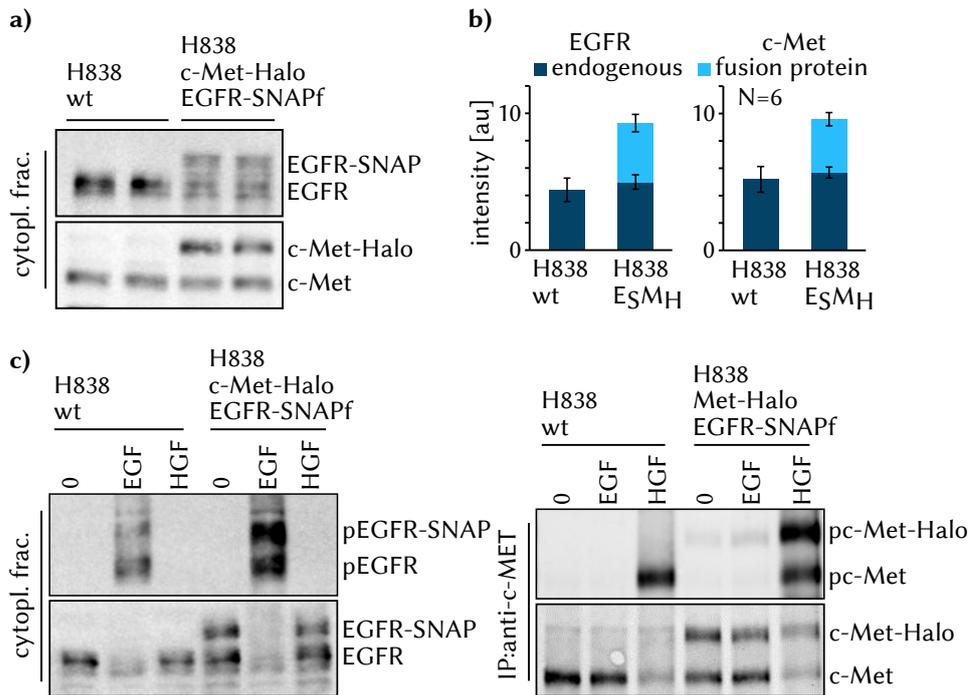


Figure 4.1.: Expression and functionality of EGFR-SNAPf and c-Met-Halo.

Total and phosphorylated protein levels of EGFR and c-Met were detected via (quantitative) immunoblotting in wildtype (wt) H838 cells and H838-EGFR-SNAPf/c-Met-Halo cells **a)** EGFR-SNAPf and c-Met-Halo was clearly detected in the cytoplasmic fraction of H838-EGFR-SNAPf/c-Met-Halo cells but not in wt H838 cells. **b)** Quantification of (a) estimated a 2.1-fold and 1.8-fold overexpression of EGFR and c-Met, respectively, with the fusion construct contributing 47% and 41% of the total protein. Error bars represent the standard deviation of six replicates. **c)** Both fusion constructs were found to be unphosphorylated in the absence of ligand (0) and phosphorylated upon stimulation with the respective ligand (EGF for EGFR-SNAPf and HGF for c-Met-Halo). The reduction of the total protein fraction upon stimulation indicated successful ligand-induced degradation for both constructs. All molecular-biological work was done by Florian Salopiata.

detected at the expected sizes of 175 kDa and 140 kDa (Fig 4.1 (a)). H838-EGFR-SNAPf/c-Met-Halo cells additionally showed the transduced EGFR-SNAPf and c-Met-Halo constructs at a size of 195 kDa and 175 kDa. Based on quantitative

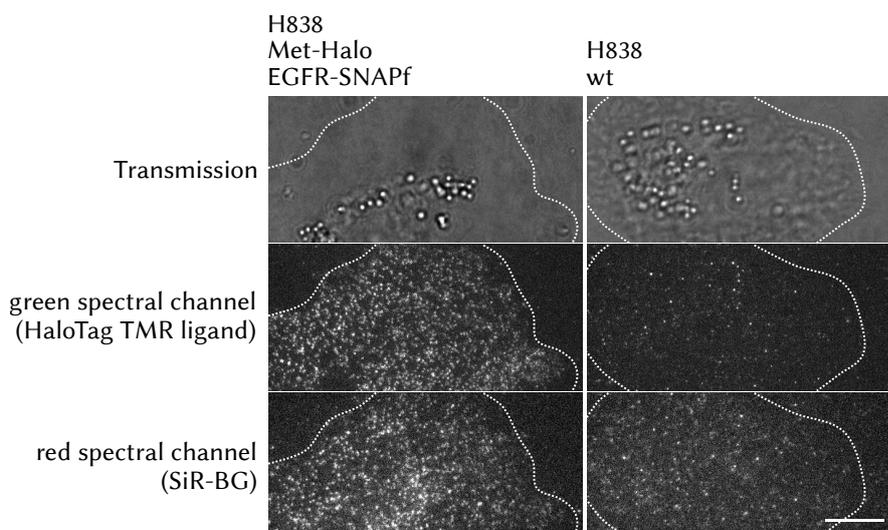


Figure 4.2.: Fluorescent labeling of EGFR-SNAPf and c-Met-Halo. Live wt H838 cells and H838-EGFR-SNAPf/c-Met-Halo cells were stained with 10 nM SiR-BG and 10 nM HaloTag TMR ligand overnight and imaged under TIRF conditions. EGFR-SNAPf and c-Met-Halo could both be specifically labeled in H838-EGFR-SNAPf/c-Met-Halo cells, while wt cells show low levels of unspecific staining. Scale bar is 10 μ m. For comparability, the intensities are consistent within each spectral channel.

immunoblotting, Florian could estimate that in H838-EGFR-SNAPf/c-Met-Halo cells, EGFR was overexpressed by a factor of 2.1 and EGFR-SNAPf contributed 47 % of total EGFR (Fig 4.1 (b)). c-Met was overexpressed by a factor of 1.8 and c-Met-Halo contributed 41 % of total c-Met.

In order to evaluate the functionality of both fusion constructs, wt cells and EGFR-SNAPf/c-Met-Halo cells were stimulated for 5 min with 40 ng/mL EGF or HGF or left unstimulated (0) and phosphorylated EGFR (pEGFR) and phosphorylated c-Met (pMet) were detected (Fig 4.1 (c)). For detection of phosphorylated and total c-Met, an anti-c-Met immunoprecipitation was performed beforehand, while for detection of phosphorylated and total EGFR the total cytoplasmic fraction was used. Neither of the fusion constructs showed a basal phosphorylation in unstimulated cells. When stimulated with their respective ligand, both constructs were effectively phosphorylated. In addition, the total protein level of endogenous receptors and fusion constructs were strongly reduced upon stimulation, indicating efficient ligand-dependent degradation.

As a further expression and functionality test of EGFR-SNAPf and c-Met-Halo, I labeled both protein tags with spectrally different dye substrates and imaged live cells on a total internal fluorescence reflection (TIRF) microscope. For this, I seeded wt H838 cells and H838-EGFR-SNAPf/c-Met-Halo cells in a cleaned Labtek and stained each cell line overnight with 10 nM siliconrhodamine benzylguanine (SiR-BG) and 10 nM HaloTag tetramethylrhodamine ligand (HaloTag TMR ligand) (See section 6.8.1 on page 121 for further details). Live-cell imaging revealed membrane-bound EGFR-SNAPf and c-Met-Halo particles (Fig 4.2). In similarly stained H838 wt cells by contrast, only a relatively small amount of unspecific staining could be detected. This experiment demonstrated that both fusions constructs were successfully expressed and transferred to the membrane, where they could be fluorescently labeled.

As a summary, EGFR-SNAPf and c-Met-Halo fusion constructs were successfully introduced into H838 cells. Both fusion constructs were moderately overexpressed fully functional in respect to localization, cell signaling and degradation. In addition, the average expression levels were found to be well suitable for single-molecule tracking.

4.2. Two-Color Single-Molecule Tracking of EGFR and c-Met

In order to directly test the hypothesis of EGFR/c-Met heterodimerization I used the two-color single-molecule tracking system described and validated in the previous section (see Fig 3.3 on page 73). In short, I simultaneously labeled EGFR-SNAPf and c-Met-Halo in H838-EGFR-SNAPf/c-Met-Halo cells with HaloTag tetramethylrhodamine ligand (HaloTag TMR ligand) and siliconrhodamine benzylguanine (SiR-BG) and imaged the so labeled cells at 37 °C on a TIRF microscope. For details on the labeling and imaging protocol see sections 6.8.1 and 6.8.2 of material and methods. To induce heterodimerization, I stimulated the cells with either 40 ng/mL EGF, 40 ng/mL HGF or 40 ng/mL of both ligands. Prior to stimulation, I imaged unstimulated cells in order to determine basal heterodimerization. For each condition, I imaged between 19–39 individual cells in three independent experiments. Fiducial makers were recorded before every experiment to generate a transformation matrix for channel registration. From all registered trajectories I extracted candidate trajectories with a spatial proximity below 1.8 pixel and a temporal overlap of at least 10 frames. I classified these candidate trajectories frame-wise as either 'dimer' or 'free', using the distance threshold derived in section 3.2.2 on page 76. As for the heterodimerization of IFNAR1 and IFNAR2, I used the GFP-enhanced staining efficiency probe (gSEP) introduced in section 2.1 on

page 2.1 as a dimer and monomer control (compare Fig 3.1 on page 70).

As a result, I could detect EGFR/c-Met heterodimers at a frequency that was significantly higher than in the monomer control in all conditions. Figure 4.3 (b) exemplifies a trajectory of an EGFR/c-Met heterodimer after stimulation with EGF and HGF. As hinted in Figure 4.3 (a) the amount of candidate trajectories with at least 5 frames of dimer state slightly increased upon stimulation with either ligand. To substantiate this, I quantified the degree of dimerization as the temporal fraction of all candidate trajectories that was classified as dimer (A candidate trajectory pair with 5 out of 10 frames classified as dimer would consequently have a degree of dimerization of 50 %). The degree of dimerization increased slightly but significantly when H838-EGFR-SNAPf/c-Met-Halo cells were stimulated with either ligand in comparison to unstimulated cells (Fig 4.3 (c)). Dimerization was promoted more strongly after stimulation with EGF than with HGF and even more for a combination of both ligand. When normalized to the dynamic range of the system set by the monomer and dimer control, the degree of dimerization increased from 20.8 % in unstimulated cells to 23.4 % after HGF stimulation, 24.3 % after EGF stimulation and 26.2 % after stimulation with both ligands. For all conditions, particles classified as dimer showed a significantly lower median jump size compared to free particles, which indicated reduced mobility of receptors in a complex (Fig A.2 (c)-(d) on page 133 of the appendix).

To quantify the ligand-dependent stability of EGFR/c-Met heterodimers, I determined the dimer-specific dissociation rate $k_{d,dimer}$ for EGFR/c-Met heterodimers from a two-component exponential fit of the distribution of dimer durations (Fig 4.4 (a)), compare equation 3.1 on page 81). As a reminder, the first component $k_{d,random}$, which was assigned to random encounters of non-interacting particles, was estimated from a one component exponential fit of the dimer durations of the monomer control and then used as fixed parameter. The second, free parameter $k_{d,dimer}$ was then compared between conditions (Fig 4.4 (b)). The dynamic range of dissociation constant based on the monomer and dimer control was found to be between 14–66 s⁻¹. When stimulated with HGF, EGF or both ligands, the dissociation rate of EGFR/c-Met heterodimers decreased from 28.1 s⁻¹ in unstimulated cells to 26.6 s⁻¹, 24.4 s⁻¹ and 23.8 s⁻¹, respectively. To better evaluate the relevance of the observed differences, I calculated the effect size *Cohen's D* between conditions in addition to significance testing. I found that the EGFR/c-Met heterodimers were significantly stabilized upon stimulation with either HGF, EGF or both ligands. When comparing the two ligands, EGF had a stronger stabilization effect than HGF ($D=1.4$ for HGF vs. $D=5.0$ for EGF) and was only slightly less potent than a combination of both ligands ($D=5.6$ for co-stimulation).

To be sure that these effects were not a consequence of an unlucky pick for

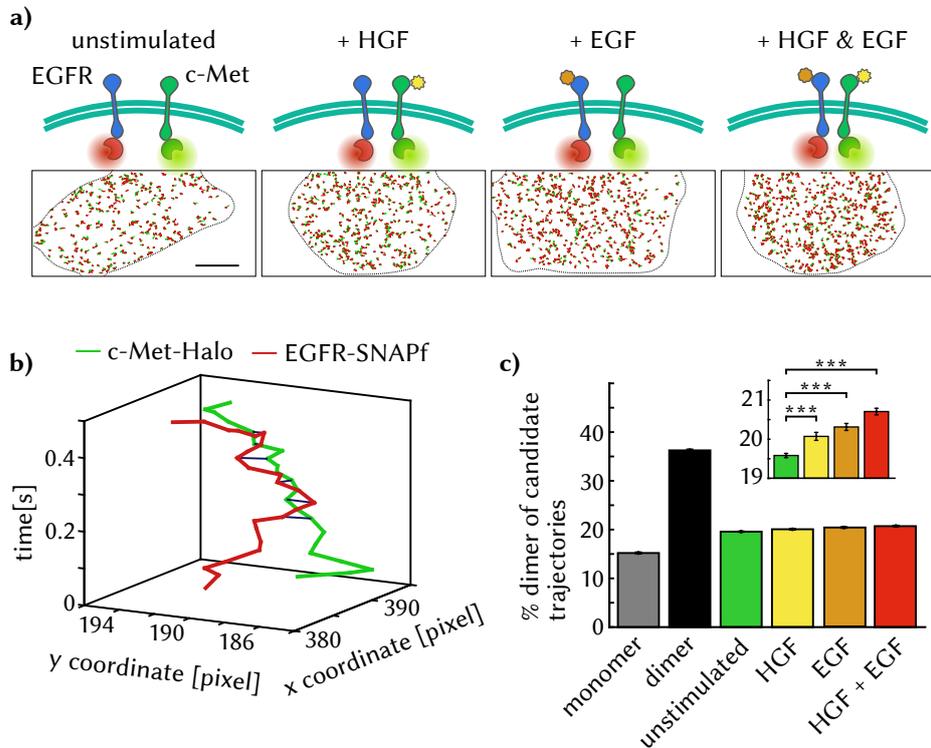


Figure 4.3.: Detection of EGFR/c-Met heterodimers. a) Overview of stimulation conditions and visualization of candidate trajectories with at least five dimer states in one example cell of each condition. H838-EGFR-SNAPf/c-Met-Halo cells were stimulated with EGF, HGF or both ligands at a concentration of 40 ng/mL. Scale bar is 10 μm **b)** Exemplified trajectory of an EGFR/c-Met heterodimer. Dimer states are symbolized as blue connections between the trajectories. **c)** temporal fraction of candidate trajectories classified as 'dimer'. Within the dynamic range of the system set by the monomer and dimer control, the degree of dimerization increased gradually from unstimulated cells to stimulation with HGF, EGF and both ligands. Errors are standard error of the mean (SEM) and significance level is $p < 0.001$ (***) based on bootstrap significance testing (see section 6.8.3 on page 123 of material and methods). The number of analyzed candidate trajectories were 7466 for the monomer control, 23 925 for the dimer control, 73 467 for unstimulated cells, 23 196 for stimulation with HGF, 31 850 for stimulation with EGF and 37 743 for stimulation with both ligands.

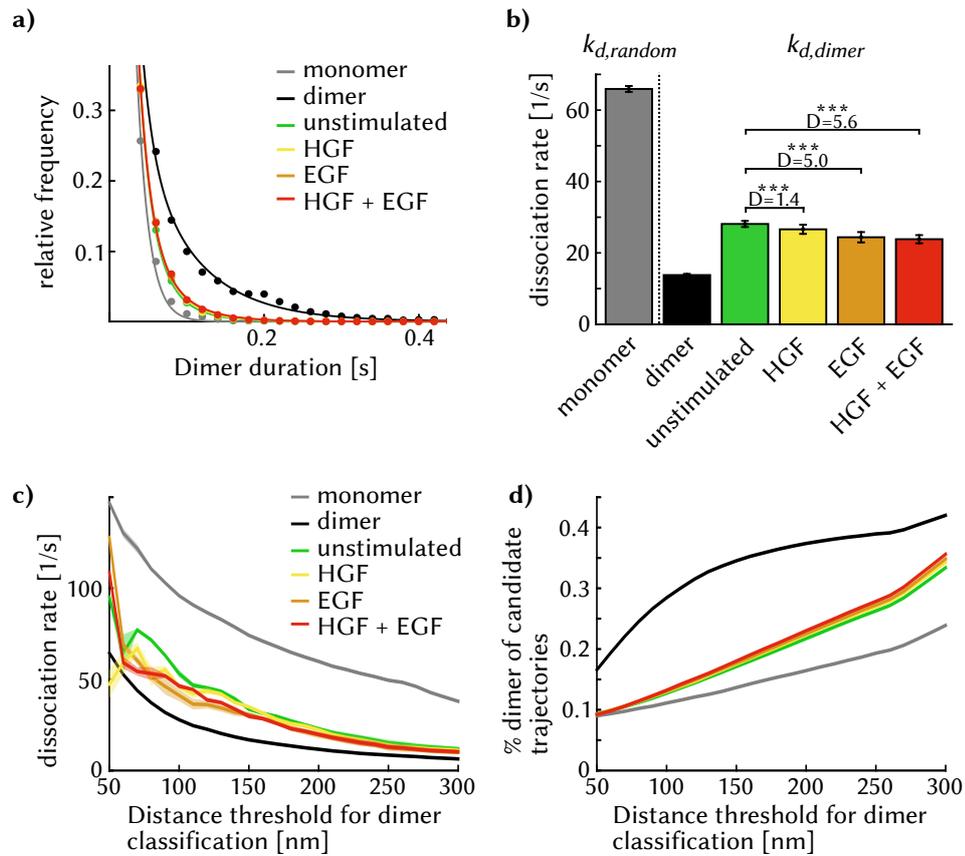


Figure 4.4.: Dissociation rates of EGFR/c-Met heterodimers and parameter variation.

a) Distribution of dimer durations for all stimulation conditions as well as the monomer and dimer control. Distributions were fitted with a two component exponential model with the first component $k_{d,random}$ being assigned to random encounters of non-interacting molecules and estimated by a one component fit of the monomer control. **b)** First or second component of the exponential fits in a). While EGFR/c-Met heterodimers were significantly stabilized upon stimulation with HGF, EGF or a combination of both ligands, EGFR had a considerably stronger stabilization effect than HGF. Errors are standard errors of the exponential model. Significance level is $p < 0.001$ (***) based on bootstrap significance testing (see section 6.8.4 on page 123 of material and methods). D represents the effect size (Cohen's D). **c)-d)** Parameter variation of the distance threshold for dimer classification. For both the dissociation rate as well as the degree of dimerization, the qualitative relation between the conditions was largely independent of the distance threshold. The absolute values of the dissociation rate and the degree of dimerization however differed approximately by a factor three and two, respectively, when the distance threshold was increased in a range of 130–300 nm

the distance threshold, I performed a parameter variation (compare page 81 of section 3.2.3). As could be seen for IFNAR1/IFNAR2 heterodimers, the quantitative dissociation rates and degrees of dimerization varied by a factor of three (from around 35 s^{-1} to 10 s^{-1}) and two (from around 15 % to 35 %), respectively, when varying the distance threshold from 130–300 nm. However, the qualitative relation between the different stimulation conditions were found to be consistent for any dimer classification distance threshold in this range (Fig 4.4 (c)-(d)).

To sum up, I could directly observe EGFR/c-Met heterodimerization using two-color single-molecule tracking and a dimerization detection based on a fixed distance threshold. In addition, I found that EGF promotes and stabilizes EGFR/c-Met heterodimers more strongly than HGF. These findings strongly support the EGFR/c-Met heterodimerization hypothesis proposed by the Klingmüller lab.

4.2.1. Heterodimerization Reduces Ligand-dependent Endocytosis of c-Met

It was suggested by Florian Salopiata that EGFR/c-Met heterodimers are partly protected from ligand-dependent endocytosis as a consequence of dimerization and should therefore show a prolonged residence time in the plasma membrane. To test this, I performed time-lapse TIRFM imaging of labeled H838-EGFR-SNAPf/c-Met-Halo cells under different stimulation conditions (Fig 4.5 (a) and section 6.9 on page 124 of material and methods). For each timepoint, Florian Salopiata determined the mean fluorescence intensity in a cellular region of interest (RIO) and estimated receptor depletion via a one-component exponential regression of the observed intensity decay. As can be seen in Fig 4.5 (b), both receptors showed a high ligand-dependent depletion when stimulated with their respective ligand (HGF for c-Met and EGF for EGFR). However, when heterodimerization was promoted by a mixture of EGF and HGF, c-Met depletion was significantly reduced. This stabilization effect was not detectable for EGFR. Because EGFR is expressed at higher levels and forms heterodimers with a variety of other receptors [54], the stabilizing EGFR/c-Met heterodimer population may have been too small to show an overall stabilizing effect on EGFR.

4.2.2. Dimer Detection Based on a Hidden Markov Model

For an independent validation of EGFR/c-Met heterodimerization, all two-color single-molecule movies and fiducial marker images were also provided to Yu Qiang of the Karl Rohr⁹ lab. He used a custom written particle detection and tracking algorithm and a fiducial marker based linear registration to generate

⁹German Cancer Research Center and Heidelberg University

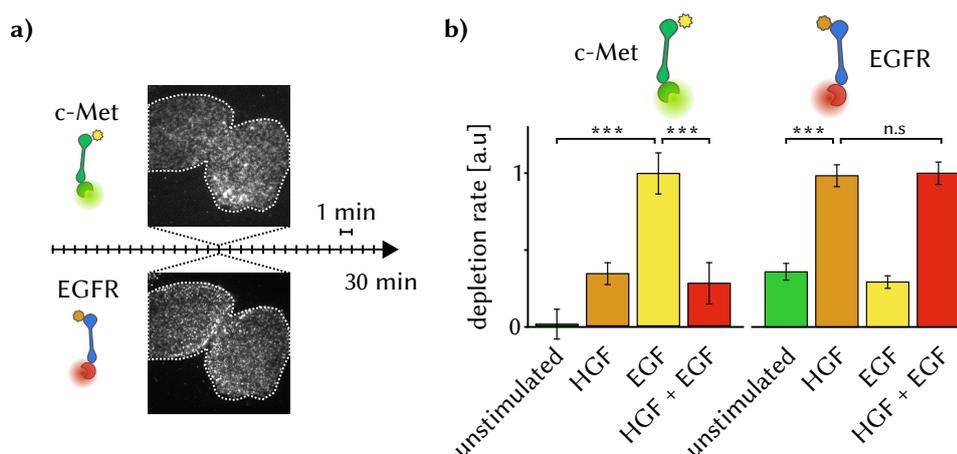


Figure 4.5.: Ligand-induced depletion of EGFR and c-Met. **a)** Scheme of the time-lapse workflow. Over a total of 30 min, c-Met-Halo and EGFR-SNAPf were imaged every minute and depletion was estimated from an exponential decay of the mean intensity within a cellular ROI. **b)** For both receptors, depletion rates significantly increased upon stimulation with their respective ligand when compared to unstimulated cells. When stimulated with a mixture of EGF and HGF, c-Met was strongly stabilized and protected from internalization. At least 10 cells were imaged per condition. Significance level is $p < 0.001$ (***) based on Student's t-test.

registered EGFR and c-Met trajectories. Instead of a fixed distance threshold, Yu applied a two-state hidden Markov model (HMM) with the particle distance as observational layer in order to classify 'free' and 'dimer' states within trajectories. To compensate for localization errors due to registration and tracking, the HMM assumed a Gaussian distance distribution between particles with a standard deviation of 3 pixel (equivalent to 312 nm). The probability to stay in the same state (Monomer→Monomer, Dimer→Dimer) was set to 0.7 and the transition probability to change states (Monomer→Dimer, Dimer→Monomer) was set to 0.3. As a further restriction, the HMM only considered EGFR or Met trajectories with a length of at least 5 frames and trajectories with a maximal jump size below 0.5 pixel were dismissed as immobile. From the durations of dimer states I then determined the dissociation rates of EGFR/c-Met heterodimers for all conditions as a one component exponential fit (Fig 4.6 (a)). The decision to use a one component exponential model instead of a two component exponential model that discretely fits unspecific and specific interactions as had been used for the interferon- α

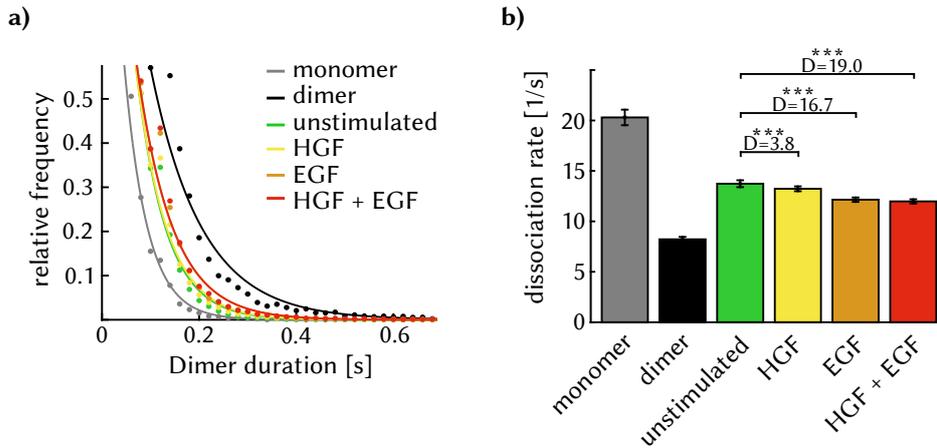


Figure 4.6.: EGFR/c-Met heterodimerization detection based on a hidden Markov model (HMM). a) Dissociation rates for every condition were estimated via a one component exponential fit of the dimer durations distributions. The yellow fit (EGF) is overlaid by the red fit (EGF+HGF) b) The resulting dissociation rates showed a significant stabilization of EGFR/c-Met heterodimers upon ligand stimulation that was stronger for EGF than HGF stimulation.

receptors, lies in the nature of the HMM. Due to the preset transition probabilities, the HMM discourages changes of states. Therefore, short lived interactions are far less likely to be classified as dimers. This feature of the HMM rejects random colocalization of non-interactive particles and largely attenuates the need for a two-component exponential model. Due to the large number of zero-bins at longer dimer duration, I performed a conversion of the standard histograms to probability densities, $P(\tau)$, as a function of event duration [81–83] using the equation

$$P(\tau_i) \approx \frac{H(\tau_i)}{0.5(\tau_{i+1} - \tau_{i-1})} \quad (4.1)$$

, where $H(\tau_i)$ is the original histogram value and τ_i is the position of bin i in an ordered list of non-zero time bins. Using this algorithm, non-zero time bins are scaled by the average time separation between a given bin and its two neighboring non-zero bins [82] and zero-bins can in return be excluded from the distribution. As a consequence, low-frequency events are mapped to a smoothly decreasing normalized probability density.

Based on the monomer and dimer control, a dynamic range of the dissociation constant between $9\text{--}21 \text{ s}^{-1}$ was found. Stimulation decreased the dissociation

rate of EGFR/c-Met heterodimers from 16.6 s^{-1} in unstimulated cells to 15.9 s^{-1} , 13.8 s^{-1} and 13.7 s^{-1} when stimulated with HGF, EGF or both ligands, respectively. This stabilization of the EGFR/c-Met heterodimers was significant for all three stimulation conditions. Again, I calculated the effect size (*Cohen's D*) between conditions and found that EGF was more potent than HGF ($D=16.7$ for EGF vs. $D=3.8$ for HGF) and nearly as potent as the combination of both ligands ($D=19.0$) in stabilizing EGFR/c-Met heterodimers compared to unstimulated cells.

While in the same order of magnitude, the HMM-derived dissociation rates of EGFR/c-Met heterodimers were approximately by a factor of 2 smaller compared to the threshold-derived dissociation rates. Again, this may be explained by the state-conserving property of the HMM. Once a trajectory is classified as dimer, it is more likely that this state is maintained as the transition probability between states was comparably low. A transient increase in particle distances will therefore not necessarily disrupt the dimer trajectory as would be the case for a fixed distance threshold. Hence, a HMM with low transition probability between states promotes dimer stability. In addition, the predefined Gaussian distance distribution between particles with a standard deviation of 3 pixel was arguably more distance-tolerant than a fixed threshold of 1.7 pixel. Considering that increasing the distance threshold was found to decrease the dissociation rate by up to a factor of 2 (compare Fig 4.4 (c)), the dissociation rates obtained by the HMM and by a fixed distance threshold were relatively comparable. Most importantly, the qualitative relations between the different stimulation conditions were independent of the analysis method. Ultimately, the HMM-based dimer analysis confirmed that EGFR/c-Met heterodimers are promoted by ligand stimulation and that a stabilization of these dimers is mostly promoted by EGF.

Part III.

Discussion and Outlook

Chapter 5.

Discussion and Outlook

A great part of cellular biology happens on short timescales, at low frequencies and in chemical equilibrium. Under these circumstances, stochastic events at very low molecule numbers may trigger positive feedback loops that eventually change cell fate. These dynamics are chronically hard to assess by classical ensemble-based methods, however their understanding is crucial to grasp the full spectrum of cell decisions in health and disease. In recent years, quantitative microscopy has evolved into a versatile field with the potential to fill this gap. Modern quantitative microscopy techniques have the temporal and spatial resolution as well as the sensitivity required to observe reactions between individual proteins and there is a clear trend towards increasing life-cell compatibility. In this dissertation, I did my part to facilitate the quantification of single-molecule fluorescence microscopy data by developing a versatile protein probe that allows to determine the degree of labeling of protein tags and fluorescent proteins. Furthermore, I established a live-cell two-color single-molecule tracking setup to quantify the dimerization of the interferon- α receptor (IFNAR) signaling complex. This dimerization is the initial trigger for the type I interferon response, which is a major barrier against viral infection. Finally, as an application with clear clinical relevance, I adapted the two-color single-molecule tracking system to examine the direct interaction between epidermal growth factor receptor (EGFR) and hepatocyte growth factor receptor (c-Met), in order to help understand what determines the sensitivity of non-small cell lung cancer (NSCLC) cells against tyrosine kinase inhibitors (TKIs) and how the formation of TKI resistance could be overcome in the treatment of NSCLC patients.

5.1. The Importance of DOL Determination for Quantitative Microscopy

Quantitative live-cell fluorescence microscopy is a highly interdisciplinary field that requires good optics, a suitable biological target molecule, a photostable

and bright fluorescent label as well as a way to specifically attach this label to the biological target. This labeling strategy eventually determines the number of fluorescent labels that is attached to each target protein and can differ from at most one (e.g. for genetically fused fluorescent proteins or protein tags) to many (e.g. for antibody labeling, especially when a sandwich of primary and secondary antibodies is used). This number of fluorescent labels per target protein is referred to as the degree of labeling (DOL) and is a crucial correction factor for the quantification of microscopy data. As mentioned in the introduction, fluorescence microscopy only observes the fluorescent labels but not the underlying target proteins. Consequently, if quantitative assumption about the target protein are to be made, the stoichiometry between target protein and its label is pivotal.

To determine the DOL of SNAPf-tag and HaloTag on a single-molecule level and within a cellular system, I have designed, cloned and expressed the GFP-enhanced staining efficiency probe (gSEP). This artificial protein allows to determine the DOL of both protein tags based on a single-molecule colocalization analysis with eGFP as an independent and practically background-free reporter. Due to the modular design, SNAPf-tag or HaloTag can readily be exchanged with any other protein tag or fluorescent protein of which the DOL should be determined. Hence, this construct represents a very versatile, yet crucial, calibration tool for quantitative microscopy. In addition to the construct itself, I have also developed an analysis algorithm that combines automated image registration of each individual cell image with a robust colocalization analysis, which I verified and characterized based on simulated single-molecule data.

I found that the DOL of SNAPf-tag and HaloTag is limited to around 0.4 and 0.5, respectively, when using the dye substrate pair HaloTag tetramethylrhodamine ligand (HaloTag TMR ligand) and siliconrhodamine benzylguanine (SiR-BG). This means that only half of the target proteins actually carried a functional label or in other words, there were twice as many proteins present as could be detected in the microscopy images. Obviously, this kind of information is absolutely crucial if absolute protein numbers or concentrations are to be determined. With the determined DOL at hand, I could estimate the absolute copy numbers of the gSEP construct on a single-cell level as an example of absolute quantification of microscopy data. Unfortunately, similarly measuring the copy number of the IFNAR, EGFR and c-Met membrane receptors used during this study was not possible in a meaningful way. Because the corresponding HaloTag and SNAPf-tag fusion constructs were over-expressed at a levels that were not reliably related to the expression levels of the endogenous receptors, no meaningful conclusions could have been drawn from the copy number of the fusion constructs about the copy numbers of the endogenous receptors. A possible solution for this problem

would be the endogenous tagging of proteins (e.g. via CRISPR/Cas9 knock-in). In such knock-in cell lines, absolute copy numbers could be determined from DOL-corrected particle densities.

In respect to the labeling of HaloTag and SNAPf-tag, I found that the highest degree of labeling can be obtained with dye substrate concentrations in the 1–50 nM range combined with prolonged incubation times of at least 3 h. Higher dye substrate concentration resulted in higher unspecific labeling rather than a higher DOL. Although not indicated by the results of this study, it cannot be excluded that labeling SNAPf-tag and HaloTag with substrate concentrations in the μM range may further increase the DOLs. Such a behavior is proposed by Liu et al. who used absorption spectrometry to determine a DOL of SNAP-tag of 0.6–0.7 after *in vitro* labeling of purified SNAP-tag protein with 10–20 μM of BG-600 for 1 h [28]. Sun et al. could obtain a DOL close to 1 for purified SNAP-tag protein after *in vitro* labeling for 1 h with 5 μM of various dye substrates and using mass spectrometry to determine the DOL [23]. These measurements however were all performed *in vitro* and on an ensemble level, where unspecific labeling is of far less concern compared to *in vivo* labeling for single-molecule microscopy. While dye substrate concentrations in the low μM range are routinely used for ensemble fluorescence microscopy methods [84, 85] and less often also for single-molecule fluorescence microscopy [86, 87], single-molecule applications generally profit from lower concentrations that avoid unspecific background. Especially when using total internal fluorescence reflection microscopy (TIRFM) to observe single molecules in the plasma membrane, my results suggest that dye substrate concentrations above 100 nM create a significant amount of unspecific background during live cell labeling. These findings are supported by Wilmes et al. who labeled the type I Interferon receptors with 30 nM HaloTag TMR ligand and 80 nM SNAP-Surface647 for an incubation time of 15 min [30] in order to perform live-cell single-molecule tracking. With these labeling conditions, Wilmes determined a DOL of SNAPf-tag and HaloTag of 0.4 and 0.3, respectively. While these values are well in line with the finding of this study, they also suggest that SNAPf-tag can be more efficiently labeled than HaloTag. The data obtained in this study rather suggests the opposite. I found that a good trade-off between specific and unspecific staining is more difficult to find for SNAPf-tag than for HaloTag, which appeared to be more readily labeled and to show less unspecific labeling. This is in line with the findings of Latty et al. who reported a DOL of HaloTag and SNAP-tag of 0.33 and 0.16, respectively, after *in vivo* labeling and cross-referencing the protein tags against an additional antibody label in fixed cells and on a single-molecule level [31]. In addition, a clear susceptibility of various SNAP-tag substrates for unspecific labeling has recently been reported by Bosch et al. [24].

The “stickiness” of some of its substrates made SNAPf-tag a suboptimal reporter

for a wider screening of labeling conditions based on single-molecule colocalization. The major hurdle is that unspecifically bound substrates can hardly be distinguished from specific labels on a single-molecule level as both types of signal appear as single emitters. This is why I introduced eGFP as a background free label into the gSEP construct. As a side result, this allowed me to quantify the maturation rate of eGFP. Because fluorescent proteins have a relatively low brightness and photostability, they are more often used in ensemble-based microscopy than on a single-molecule level. This might be why the maturation rate is in the literature mostly understood as the kinetic of fluorescence protein folding (how fast do fluorescent proteins fold) and not its efficiency (how often is the folding successful). Indeed, relatively few people have asked this question and the reported fractions of functional, fluorescent proteins vary in the literature. For eGFP, maturation rates between 0.7–1 [88, 89] have been reported. For different fluorescent proteins, a high inter-protein difference of the maturation rate can be recognized with maturation rates ranging from 0.85 (YFP) [90] to 0.4 (mCherry) [89] or even as low as 0.2 (mRFP) [89]. These findings make it less surprising that the eGFP protein used in this study was found to have a low maturation rate of just 0.29. Probably the most important lesson from this is the fact that fluorescent proteins may routinely have a maturation rate below one. So they too, require a proper determination of their fluorescent functional fraction when used for quantitative microscopy.

A remaining question concerning the determination of the DOL and maturation rate reported in this study is its generalizability. The reaction kinetic between a protein tag and its dye substrate is certainly influenced by the chemical properties of the substrate. When labeling an intracellular target, the effective concentration of the dye substrate is additionally limited by its cell permeability, which depends foremost on the substrate itself but may also be cell-type dependent. Similarly, the folding efficiency of the protein tags or fluorescent proteins may be different for different cell types. To target these questions, the DOL of HaloTag and SNAPf-tag as well as the maturation rate of eGFP is being determined for different dye substrates and in different cell lines by Klaus Yserentant, Felix Braun and Wioleta Chmielewicz of the Herten group.

As a major result, I could show that neither fluorescent proteins nor protein tags can guarantee a stoichiometric 1:1 ratio between target protein and fluorescent label. In both cases, a proper determination of the DOL is required to generate true quantitative data. Hence, this findings encourage a proper DOL determination for any labeling strategy and any target system and propose a calibration probe in the design of the gSEP construct as a way to do so *in situ* and on a single-molecule level.

5.2. Development of a Two-Color Single-Molecule Tracking System

As mentioned, fluorescence microscopy techniques need to combine a high temporal and spatial resolution with single-molecule sensitivity in order to assess transient and rare protein-protein interactions. One such technique is single-molecule tracking. In this study, I implemented a two-color single-molecule tracking system that allowed to observe and quantify interactions between membrane receptors of live cells in real time. To benchmark and validate this system, I made use of simulated single-molecule data and reused the above mentioned gSEP construct, however this time as an artificial membrane-associated protein that could model strict monomers or strict dimers, depending on the labeling strategy. On the software side, I developed an automated linear image registration and a distance-based classification algorithm to detect and quantify dimerization events.

As a proof of principle, I could well discriminate between the monomer and dimer control in respect to dimer frequency and stability. I could further confirm the ligand-induced heterodimerization between Interferon- α receptor 1 (IFNAR1) and Interferon- α receptor 2 (IFNAR2) in Huh7.5 cells. Finally, I could directly confirm EGFR/c-Met heterodimers as a proposed regulator of lung cancer sensitivity towards tyrosine kinase inhibitors.

More generally, I could confirm that multi-color single-molecule tracking is a powerful method to observe protein-protein interactions on low timescales and frequencies and with little disturbance of the underlying biological system. I proofed the usefulness of this technique both for basic biophysical research as well as for clinically relevant applications in biomedical research.

5.2.1. Comparison of the Obtained IFNAR Dimerization Dynamic with Literature

The heterodimerization of IFNAR1 and IFNAR2 upon ligand binding is a well described initiator of the cellular type I interferon response. Its dynamic has been described on a single-molecule level and in live cells by Stephan Wilmes [30] and most recently by Changjiang You [91] from the group of Jacob Piehler¹. In general, they reported slower diffusion of the receptors ($0.23 \mu\text{m}^2 \text{s}^{-1}$ in this study vs. $0.09 \mu\text{m}^2 \text{s}^{-1}$ reported by Wilmes and $0.06 \mu\text{m}^2 \text{s}^{-1}$ reported by You). Moreover, the dissociation rate of IFNAR1/IFNAR2 heterodimers determined by You was more than 2 orders of magnitude smaller than the dissociation rate found in this study (23s^{-1} in this study vs. 0.07s^{-1} reported by You). There are several factors

¹Osnabrück University

that may partly explain these differences and help to harmonize both results.

First, Wilmes and Yu both labeled the IFNAR receptors with a sandwich labeling strategy where they fused the receptor chains with SNAPf-tag and HaloTag, respectively, and then subsequently bound biotin, streptavidin and biotinylated quantum dots to the receptors. It can be expected that the large molecular weight of these labels, which are approximately 20 times larger than using only protein tags and organic dye substrates, as well as the larger hydrodynamic radius may slow receptor dynamics to some degree.

Second, while I performed all tracking experiment at 37 °C, Wilmes and You decided to measure receptor kinetics at room temperature in order to reduce basal internalization of the receptors and thus prolong receptor trajectories [30]. Reducing the temperature however, generally slows diffusion and reaction kinetics. With special regard to the dissociation rate, a lower temperature provides less thermal energy that may counteract the receptor binding energy. Especially when receptors bind weakly, lowering the temperature may largely stabilize these complexes. The exact impact on binding kinetics and diffusion however is hard to estimate. The temperature dependence of membrane protein diffusion has been reported to be both protein and cell type dependent [92, 93]. For different GPI-anchored reporter molecules, lowering the temperature from 37 °C to room temperature has been shown to reduce the diffusion coefficient in a range between 30–60 % [92]. For some membrane anchored proteins, diffusion was reported to be mostly temperature independent [93]. For such temperature-independent reporters, the interaction with the cytoskeleton has been supposed to govern diffusion rather than temperature. As an additional parameter, the higher temporal resolution of 20 ms in this study compared to 32 ms used by Wilmes and You may have further increased the measured diffusion coefficient as an effect of a higher sampling of predominantly random motion.

To estimate the influence of temperature and frame rate on IFNAR diffusion, I recorded some single-particle tracking movies of IFNAR1 at room temperature with a frame rate of 20 Hz (50 ms exposure time), at 37 °C with a frame rate of 20 Hz and at 37 °C with a frame rate of 50 Hz (20 ms exposure time). As mentioned, the average diffusion coefficient of IFNAR1 at 37 °C and 50 Hz frame rate was found to be $0.23 \mu\text{m}^2 \text{s}^{-1}$ (based on 53 941 trajectories). Reducing the frame rate to 20 Hz lowered the average diffusion coefficient of $0.13 \mu\text{m}^2 \text{s}^{-1}$ (based on 8485 trajectories). Additionally lowering the temperature to room temperature further reduced the diffusion coefficient by 54 % to around $0.06 \mu\text{m}^2 \text{s}^{-1}$ (based on 15 042 trajectories). Hence, when adapting the experimental conditions, the diffusion properties reported by Wilmes and You could be well reproduced. It may be assumed however, that the higher frame rate and temperature used during this study is more suitable to reveal the physiological dynamics of the IFNAR receptors.

On the other hand, these experimental differences can probably not explain the two order of magnitude difference in the observed dissociation rates. As already mentioned, Wilmes and Yu used quantum dots as fluorescent labels. One disadvantage of quantum dots as fluorescent labels is their relatively strong fluorescence blinking, i.e. a transition between a fluorescent 'on' and 'off' state. These 'off' states easily lead to broken trajectories as particles cannot be detected for one or several frames of the single-molecule tracking movie. To compensate for this, both Wilmes and You introduced pre-analysis steps that favored long dimer trajectories. While Wilmes rejected any dimer trajectories shorter than 10 frames (0.3 s) from further analysis, You performed a time-lapse particle correlation with a time window of 20 frames (0.6 s) to create longer dimer trajectories from shorter, fragmented trajectories within a diffusion-limited area.

As an advantage compared to organic dyes, quantum dots are markedly more photostable, so Wilmes and Yu could record longer trajectories of up to several minutes duration. Due to the faster photobleaching of organic dyes, the trajectories in this study were limited to a duration of several seconds. As the lifetime of the labels limits the observation window, the duration of heterodimerization events is restricted to several seconds, too. However, the observed dimerization events in this study were not found to reach this limitation but showed a lifetime that was still significantly shorter than the dimer lifetime of the dimer control, which was used to benchmark the dynamic range of the dimer detection.

While all these factors help explain the shorter dimer lifetimes found in this study when compared to the work of You and Wilmes, it is hard to conclude if the observed differences come down to experimental differences alone or if a different physiology could be observed. Not to forget that Wilmes and You also used a different cell line. Further experiments with a stepwise accommodation of the experimental parameters would be necessary to see if the quantitative dynamic of IFNAR interaction found in the three studies can be aligned. Qualitatively however, both studies agree on the ligand induced heterodimerization of IFNAR1 and IFNAR2 that is accompanied by a reduced mobility of the dimers.

As a previously only theoretically assumed property of ternary signaling complexes, I could experimentally confirm that ligand oversaturation can lead to a reduced complex formation. While at physiological ligand concentrations IFNAR1 and IFNAR2 share and are bound by one molecule of IFN- α or IFN- β , a very high ligand concentration increases the possibility that each subchain binds an individual ligand molecule. In this situation, dimerization is reduced as the ligand loses its function as cross-linking molecule.

5.3. Implications of EGFR/c-Met Dimerization for Lung Cancer Therapy

As described in the introduction, a major hurdle of treating late-stage lung cancer patients with tyrosine kinase inhibitors (TKIs) is the emergence of TKI resistance. In many patients, TKI resistance coincides with an amplification of the MET gene and a resulting overexpression of the hepatocyte growth factor receptor (c-Met). While crosstalk between the EGFR and c-Met signaling pathways is believed to govern TKI sensitivity, the exact mechanism by which c-Met overexpression provides TKI resistances is poorly understood. To understand the relationship between growth hormone signaling, cell growth and TKI sensitivity in the context of non-small cell lung cancer (NSCLC), the groups of Ursula Klingmüller² and Jens Timmer³ have developed a dynamic pathway model of combined EGFR and c-Met signaling based on time resolved quantitative western blot data. This model and the underlying data suggested that EGFR/c-Met heterodimerization crucially modulates EGFR and c-Met signaling as well as the sensitivity of NSCLC cells against TKIs, especially when c-Met is overexpressed.

As a general idea, tumor progression is promoted by growth hormone signaling through EGFR and c-Met (Fig. 5.1 (a)). By blocking EGFR signaling, tyrosine kinase inhibitors effectively reduce tumor growth. This effect can be overcome by MET amplification, as overexpressed c-Met compensates for blocked EGFR signaling in driving cell growth and tumor progression.

According to the dynamic pathway model, the formation of heterodimers between EGFR and c-Met may even enhance this effect by reducing the ligand-dependent degradation of heterodimers and thereby prolonging c-Met signaling. Using time-lapse live-cell fluorescence microscopy, I could directly prove this enhanced surface stability of c-Met under conditions that promote heterodimerization with EGFR.

While this increased surface stability presents EGFR/c-Met heterodimerization as a further driver of tumor progression, the direct interaction with EGFR was likewise predicted to make c-Met vulnerable to inhibition by TKIs. Hence, by forming heterodimers with EGFR, c-Met can indirectly be targeted by TKIs, which makes the degree of EGFR/c-Met dimerization a valuable predictor for the efficiency of TKI treatment.

The likelihood that c-Met forms a heterodimer with EGFR depends primarily on the number of available EGFR receptors. Consequently, a low c-Met/EGFR ratio favors c-Met heterodimerization with EGFR, while a high c-Met/EGFR ratio

²German Cancer research Center, Heidelberg

³Freiburg University

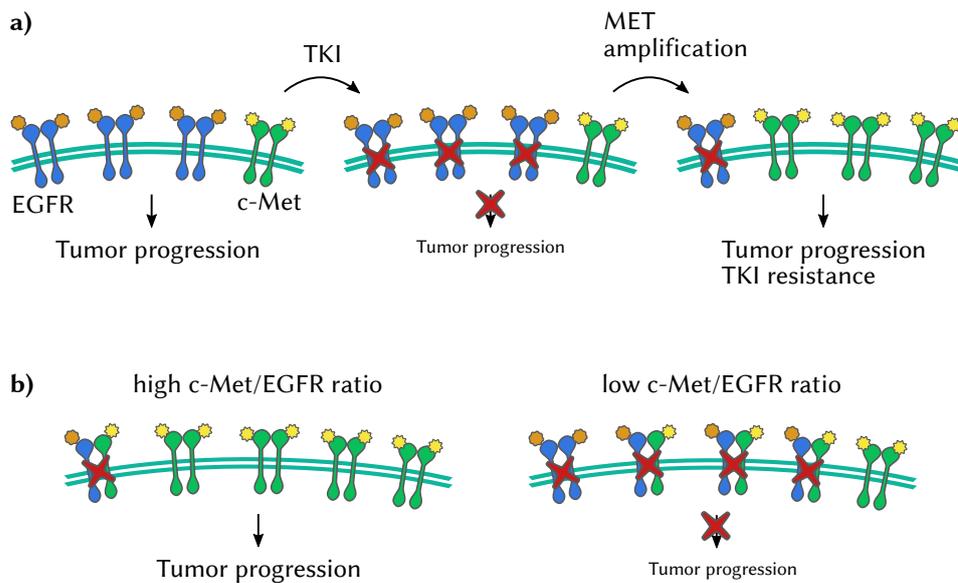


Figure 5.1.: Influence of EGFR and c-Met expression levels on tumor progression and TKI sensitivity. a) EGFR and c-Met signaling promotes tumor progression through partly overlapping pathways. When EGFR signaling is inhibited by tyrosine kinase inhibitors (TKIs), MET amplification and the associated overexpression of c-Met further drive tumor growth. **b)** By forming a heterodimer with EGFR, c-Met can be indirectly inhibited by TKIs. Governing the efficiency of EGFR/c-Met heterodimerization, the ratio between the expression levels of EGFR and c-Met may serve as a valuable biomarker for TKI responsiveness.

promotes the formation of TKI-insensitive c-Met homodimers (Fig. 5.1 (b)). So rather than the absolute copy number of c-Met, the expression ratio between c-Met and EGFR determines TKI sensitivity and even patients with a MET amplification may respond well to TKIs if their tumor cells show a similarly high expression of EGFR. Moreover, patients with a low c-Met/EGFR ratio will likely profit from assisting therapies that reduce the expression of c-Met, e.g. the bivalent c-Met-targeting antibody MM-131 developed by Merrimack Pharmaceuticals. Hence, under the assumption of EGFR/c-Met heterodimerization, the c-Met/EGFR ratio can be proposed as a biological marker to guide TKI treatment and possible application of c-Met-targeting antibodies.

For a direct proof of this model, I could observe heterodimerization between EGFR and c-Met using two-color single-molecule tracking. As an additional verification of model predictions, I could show that EGFR/c-Met heterodimerization

was mainly promoted and stabilized by stimulation with epidermal growth factor (EGF). This finding was verified by an independent analysis based on a hidden Markov model performed by Yu Qiao of the group of Karl Rohr⁴.

Generally, the easy adaption of the two-color single-molecule tracking routine from the IFNAR to the EGFR/c-Met system showed its versatility and its general potential to unravel rare receptor interaction dynamics on fast timescales. It should be well adaptable to any further receptor pair.

A fully comprehensive quantification of receptor numbers and dynamics in this study was hampered by the poorly-defined relation between the expression level of the (fluorescent) receptor fusion constructs and the (dark) endogenous receptors. While I showed based on the gSEP construct that absolute protein number are accessible, the absolute numbers of the receptor fusion constructs will not provide meaningful information about the endogenous receptors. In the advent of the CRISPR/Cas9 genome-editing technique, direct knock-ins of fluorescent reporters are becoming more feasible. Circumventing the problems of overexpression and untagged endogenous receptors, such modifications will certainly help unlock the full potential of quantitative microscopy and determine protein-protein interactions in a well-defined, yet physiological, environment.

⁴Heidelberg University

Part IV.

Materials and Methods

Chapter 6.

Methods and Reagents

6.1. Molecular Cloning

6.1.1. Polymerase Chain Reaction

All polymerase chain reactions (PCRs) were performed using Phusion Hot Start II DNA Polymerase (Biozym) in Phusion HF-Buffer and according to manufacturer's instruction [94]. In a total reaction volume of 50 μ L per reaction, 10 ng of template DNA, 200 μ M deoxyribonucleotide triphosphates (dNTP), 1 unit of Phusion Hot Start II DNA Polymerase and 500 nM of each primer was used. To equip a PCR product with restriction sites, primers that contained the respective restriction site as overlap were used. While the manufacturer's recommended cycling conditions were used, the primer annealing temperature was predicted using the software Geneious v5.3.6 (Biomatters) and the duration of the elongation step was adjusted to the size of the PCR product assuming that Phusion Hot Start II DNA polymerase produces 2 kbp per min.

6.1.2. Restriction Digest

For the digestion of plasmids and PCR products that were used for further cloning steps, 1 μ g of DNA was digested for 2 h at 37 °C in a 50 μ L reaction using 20 units of each reaction enzyme in the recommended buffer. All restriction enzymes and buffers were purchased from New England BioLabs (NEB) and used according to manufacturer's instructions [95].

To validate the composition of plasmids, a slightly reduced test digestion was performed in a total volume of 20 μ L and only 10 units of each reaction enzyme. The amount of DNA was kept at 1 μ g per digestion reaction to assure a good visibility of the digestion products on an agarose gel.

6.1.3. DNA Purification

PCR products were purified using QIAquick PCR Purification Kit (Qiagen, Cat No. 28106) according to manufacturer's instructions [96]. In short, PCR products were bound to a centrifugation column, washed with an ethanol-based buffer and eluted with 10 mM TrisHCl, pH 8.5. While the column binds DNA fragments with a size between 100 bp and 10 kbp, proteins and small DNA or RNA fragments like PCR primers are removed during the washing step. The QIAquick PCR Purification Kit was also used for the purification of digestion products, when small DNA fragments below 20 bp, e.g. sticky ends of PCR products or very small inserts, had to be removed.

For the purification and separation of larger digestion products, gel extraction was performed. For this, the digestion reaction was loaded on a 1 % agarose gel (UltraPure Agarose, Invitrogen) in 0.5x TAE-Buffer¹ and 1:20000 MidoriGreen (NIPPON Genetics Europe GmbH) and digestion products were separated for 40 min at 80 V. The desired DNA fragments were excised from the gel and extracted using QIAEX II Gel Extraction Kit (Qiagen, Cat. No 20051) according to manufacturer's instructions [97]. In short, the gel was dissolved and the DNA fragments were bound to beads, washed with an ethanol-containing buffer and finally eluted.

6.1.4. Ligation and Transformation

Purified digestion products were ligated using Quick T4 DNA Ligase (New England Biolabs) in a 20 µL reaction for 10 min at room temperature according to manufacturer's instructions [98]. Typically, one ligation reaction with a equimolar vector-to-insert ratio and an additional reaction with a three-fold excess of insert were set up. As a negative control, one ligation reaction that contained no insert was prepared.

For the transformation of competent *e.coli* DH5- α (Subcloning Efficiency, Invitrogen), 50 µL of *e.coli* DH5- α per transformation reaction were thawed on ice and 5 µL of the ligation reaction was added. The uptake of the plasmid was promoted by heat shock. For this, the bacteria were kept on ice for another 25 min, warmed to 37 °C for 5 min in order to permeabilize the membrane and cooled on ice for another 10 min. The transformed bacteria were grown for 30 min at 37 °C in 1 mL of S.O.C medium (Invitrogen, Cat. No 1554034) and subsequently centrifuged for 1 min at 5000 rpm. The volume of the supernatant was reduced to 100 µL, the bacteria were resuspended and streaked on a lysogeny broth (LB)-agar plate

¹0.5x TAE Buffer: 242 g Tris, 100 mL 0.5 M EDTA (Dihydrat, Sigma), in 50 L, pH 8.0

containing $100 \mu\text{g mL}^{-1}$ ampicillin. The plates were then incubated overnight at 37°C .

6.1.5. MiniPrep

For small-scale amplification of plasmids, single colonies of overnight grown, transformed *e.coli* DH5- α were picked from the LB-agar plate and grown overnight in 3 mL LB medium containing $100 \mu\text{g mL}^{-1}$ ampicillin. QIAprep Spin Miniprep kit (Qiagen, Cat No. 27106) was used according to manufacturers instructions [99] to isolate plasmid DNA. In short, 1.5 mL of this culture were lysed and plasmid DNA was purified via a DNA-binding column. To identify *e.coli* DH5- α colonies that express the correct plasmid, test-digestions of the purified plasmid DNA were performed.

6.1.6. MaxiPrep

For large-scale amplification of plasmids, 100 μL of transformed *e.coli* DH5- α overnight culture was used to inoculate 100 mL of LB-medium containing the selection antibiotic. After incubation at 37°C overnight, plasmid DNA was extracted using either Jetstar 2.0 MaxiPrep Kit (Genomed) or PureLink HiPure Plasmid Maxiprep Kit (Invitrogen) following manufacturers' instructions [100, 101]. Isolated plasmid DNA was solved in double-distilled, sterile water at a final concentration of $1 \mu\text{g } \mu\text{L}^{-1}$. Plasmid concentration was measured using a Nanodrop UV-Vis spectrometer (Thermo Fisher Scientific) and adjusted to $1 \mu\text{g } \mu\text{L}^{-1}$ with double-distilled, sterile water. To verify the isolated plasmid, the insert region was sequenced from both sides by Eurofins Scientific.

6.2. Cell Culture

6.2.1. Retroviral Transduction

Stable transgenic human cell lines were generated using the third generation of the Phoenix Amphotrophic retroviral system developed in the Nolan lab at Stanford University [102]. The system uses a modified Moloney Murine Leukemia Virus (MMULV) to deliver transgenic DNA to dividing, mammalian cells. For its replication, the virus depends on a 293T cell line-derived producer cell line called Phoenix. The viral factors necessary for viral replication are spatially separated and encoded partly on the Phoenix genome and partly on the retroviral cloning vectors pMOWS or pBABE. Hence, viral particles can only form when Phoenix cells are transfected with one of these retroviral vectors. The resulting viral particles carry

only the genetic information encoded on the retroviral plasmid. So although they can infect mammalian cells, they can not further replicate in the absence of Phoenix cells. Due to the infectivity of the viral particles, all steps of viral transduction were carried out under biosafety level 2.

To produce viral particles, Phoenix amphi cells were thawed and cultivated in Dulbecco's Modified Eagle Medium (DMEM, Gibco Cat No. 31 053 028) supplemented with 10 % fetal bovine serum (FBS, Gibco Cat No. 10 500 064), 1 mM sodium pyruvate (Gibco Cat No. 11 360 039), 2 mM L-glutamine (Gibco Cat No. 25 030 149) and 50 U mL⁻¹ Penicillin-Streptomycin (Pen/Strep, Invitrogen) for at least 2 passages. 800 000 Phoenix amphi cells were then seeded in a six-well plate and grown overnight. The next day, medium in all wells was changed to DMEM containing 25 µM chloroquine. Per well, 8 µg of plasmid DNA was diluted in 112.5 µL double-distilled water in a 15 mL tube and 125 µL of 2.5 M calcium chloride solution was added. To precipitate the DNA, 125 µL of 2x HEPES buffered saline² (HBS) was drop-wise added while vortexing and the solution was left at room temperature for 1 min. This solution was drop-wise added to the Phoenix cells. After 6–8 h incubation at 37 °C and 5 % CO₂, medium was changed to DMEM (supplemented with 10 % FBS and 50 U mL⁻¹ Pen/Strep) and cells were incubated overnight at 37 °C and 5 % CO₂. 16–18 h after transfection, the supernatant was harvested and filtered through a 0.45 µm Millex-HA Millipore filter (Merck). The supernatant contains infectious viral particles and the filtration step ensures that no Phoenix cells are propagated together with the virus. This supernatant was stored at –80 °C or used immediately to infect the target cells.

Target cells (usually Huh7.5) were seeded one day before infection in a six well plate at a density of 100 000 cells/well. 250 µL of virus supernatant were mixed with 750 µL DMEM and 1 µL of 8 mg mL⁻¹ polybrene (Hexadimethrinbromid, Sigma). For each well of the six well plate, the medium was removed and 1 mL of the virus supernatant dilution was added to the cells. Cells were then spin-infected for 3 h at 340 g and room temperature. After spin-infection, medium was changed to DMEM (with 10 % FBS and 50 U mL⁻¹ Pen/Strep) and cells were incubated overnight at 37 °C and 5 % CO₂. The next day, selection with the respective antibiotic was started (puromycin at a concentration of 1.5 µg mL⁻¹ for single-transduced cell lines or a combination of 0.75 µg mL⁻¹ puromycin and 200 µg mL⁻¹ G418 for co-transduced cells). After three passages, cells could be declared virus-free. They were subsequently frozen or transferred to biosafety level 1 for further culture and experiments.

²2x HBS: 1.64 g NaCl, 1.19 g HEPES, 0.021 g Na₂HPO₄ in 100 mL double-distilled water adjusted to pH 7.05 with NaOH

6.2.2. Splitting, Freezing and Thawing of Cells

All cells were cultivated in DMEM supplemented with 10 % FBS, 1 mM sodium pyruvate and 2 mM L-glutamine or Glutamax (Gibco Cat No. 35 050 061). This will be referred to as growth medium. Depending on the resistance gene, 1.5 $\mu\text{g mL}^{-1}$ puromycin (Sigma) or 400 $\mu\text{g mL}^{-1}$ G418 (Sigma) were added to maintain stably transduced cell lines. Co-transduced cell lines that contain both resistance genes were maintained using both antibiotics at half the concentration mentioned above. Cells were grown in a 10 cm tissue culture dish (TPP) at 37 °C and 5 % CO₂. Every two to four days or when cells reached a confluence above 90 %, two to three million cells were seeded in a new culture dish. For this, the near confluent cells were washed with 5 mL Dulbecco's phosphate buffered saline³ (DPBS, Sigma Cat No. D8537) and incubated with 2 mL TrypLE Express (Gibco Cat No. 12 604 021) for 5 min at 37 °C and 5 % CO₂. Cells were rinsed and resuspended with 8 mL growth medium to yield a total volume of 10 mL. 1–4 mL of this cell suspension was transferred to a new cell culture dish and growth medium was added to a total volume of 10 mL.

For long-term storage in liquid nitrogen, cells were grown in a 15 cm tissue culture dish (TPP) to 95 % confluence. Cells were then detached as described above, pelleted for 2 min at 1000 rpm and resuspended in 8 mL freezing medium, which contains 90 % FBS and 10 % dimethyl sulfoxide (DMSO, Sigma). Cells were transferred to 2 mL cryo tubes (Greiner) at 1 mL per tube. All cryo tubes were cooled to –80 °C in isopropanol overnight and transferred to liquid nitrogen.

To thaw cells from liquid nitrogen, frozen cells were thawed at 37 °C and immediately transferred to a 10 cm tissue culture dish with an additional 9 mL prewarmed growth medium. After incubation overnight, medium was replaced by fresh growth medium containing the respective selection antibiotics if applicable.

6.3. Coupling of Siliconrhodamine Benzylguanine

50 μL of 2 $\mu\text{g mL}^{-1}$ silicon rhodamine N-hydroxysuccinimide ester (SiR NHS ester, Spirochrome) were coupled to a three-fold molar excess of O6-benzylguanine amine (BG, synthesized by Dr. Dominik Brox⁴) in 50 μL anhydrous dimethylformamid (DMF) and 10 μL N,N-diisopropylethylamine (Dipea). The reaction mixture was incubated for 3 h at 40 °C in a shaker. Siliconrhodamine benzylguanine (SiR-BG) as the reaction product was purified via high performance liquid chromatography (HPLC) and dried using a rotary evaporator. SiR-BG was solved

³DPBS: 0.2 g L⁻¹ KCl, 0.2 g L⁻¹ KH₂PO₄, 8 g L⁻¹ NaCl, 1.15 g L⁻¹ Na₂HPO₄ (anhydrous)

⁴AK Herten, Heidelberg University

in DMF at a final concentration of 120 μM , as measured using an absorption spectrometer. The identity and purity of SiR-BG was confirmed via mass spectrometry. Here, the expected mass of 747.28 U for sodium-complexed SiR-BG was found. For labeling experiments, the desired amount of the stock solution of SiR-BG in DMF was diluted in growth medium to yield a staining stock solution (typically 1 μM SiR-BG in DMEM).

6.4. Cleaning of Labteks

In order to allow for an efficient total internal reflection for TIRF microscopy, cells were seeded in glass-bottom, 8-well cell chamber (Nunc Lab-Tek chambered coverglass, Thermo Fisher Scientific). Before seeding the cells, these chambers had to be thoroughly cleaned to reduce background on the single-molecule level. For this, all wells of the chamber were filled with 250 μL of 0.1 M hydrofluoric acid in ultrapure water, incubated for 1 min at room temperature and washed twice with DPBS. This procedure was repeated two more times. After the final washing step, the chambers were filled with DPBS and the chamber was stored at 5 $^{\circ}\text{C}$ until further use but not longer than three days. Usually, LabTeks were cleaned right before cells were seeded.

6.5. Single-Molecule Alternating Laser Excitation (smALEX) Measurements

Huh7.5-pBABE-gSEP cells were seeded in six wells of a cleaned LabTek. Two wells were left unstained, two wells were stained overnight with 1 nM HaloTag tetramethylrhodamine ligand (HaloTag TMR ligand, Promega) alone and the two remaining wells were stained overnight with 1 nM HaloTag TMR ligand and 1 nM SiR-BG. The next day, all cells were washed three times with full growth medium over one hour, fixed with 4 % paraformaldehyde (PFA, Sigma) for 40 min and washed two more times with growth medium. The so prepared cells were imaged at room temperature on a Nikon TiE TIRF microscope. Acceptor and donor emission was recorded synchronously using a spectral image splitter (Optosplit III, Cairn Research) to direct the donor signal to the upper half of the camera chip, and the acceptor signal to the lower half. The sample was excited with alternating laser excitation at 20 Hz, triggered by the frame rate of the camera. Hence, donor and acceptor were sequentially excited for one frame each. To determine FRET between eGFP and HaloTag TMR ligand, excitation was alternated between 1.8 mW at 488 nm and 1.6 mW at 561 nm (power of the laser beam as it leaves the objectives) and the emission light was cleaned up using a 525/50 bandpass filter for eGFP

and a 605/70 bandpass filter for TMR emission (both AHF Analysetechnik). FRET between Halo-TMR and SiR-SNAP was determined using alternating excitation with 1.6 mW at 561 nm and 4.8 mW at 640 nm, while the emission light was cleaned up using a 605/70 bandpass filter for TMR and a 685/50 bandpass filter for SiR emission (both AHF Analysetechnik). A total of 200 frames were recorded, which equals 100 ALEX-frames. For each pair of fluorophores, a short video of multicolor beads (TetraSpeck Microspheres 0.2 μm , Thermo Fisher Scientific) was recorded and used as fiducial marker calibration to align the two spectral channels. FRET between the two fluorophore pairs was determined using the Matlab-based software ALEX that was developed and provided by Dr. Kristin Größmayer⁵. For each pair of channels, a non-linear registration matrix was calculated based on at least 40 manually selected landmarks in the fiducial marker images. Based on this image registration, particles were detected in both channels using a two-dimensional Gaussian intensity fit with a manually set intensity threshold. Particle pairs with a distance below one pixel were classified as colocalized. For each particle, the intensity of the particle itself and the intensity of a surrounding ring with a width of 3 pixels was determined in order to estimate background. FRET efficiencies E and stoichiometries S were calculated frame-wise based on the background corrected intensities of each pair of colocalized particles for each ALEX frame using the following equations:

$$\begin{aligned}
 E &= \frac{F_{D_{ex}}^{Aem}}{F_{D_{ex}}^{Aem} + F_{D_{ex}}^{Dem}} \\
 S &= \frac{F_{D_{ex}}^{Aem} + F_{D_{ex}}^{Dem}}{F_{D_{ex}}^{Aem} + F_{D_{ex}}^{Dem} + F_{A_{ex}}^{Aem}}
 \end{aligned}
 \tag{6.1}$$

, where E and S are FRET efficiency and stoichiometry, respectively, $F_{D_{ex}}^{Aem}$ is the fluorescence intensity of the acceptor fluorophore under donor excitation, $F_{D_{ex}}^{Dem}$ is the fluorescence intensity of the donor fluorophore under donor excitation and $F_{A_{ex}}^{Aem}$ is the fluorescence intensity of the acceptor fluorophore under acceptor excitation.

Direct excitation was corrected for using a excitation coefficient of $D = 0.124$ for the direct excitation of TMR at 488 nm [103] and a excitation coefficient of $D = 0.056$ for the direct excitation of SiR at 561 nm [104]. Spectral bleed-through was not corrected for. For calculating the average FRET efficiency and stoichiometry only E and S values within an interval of -0.25 to 1.25 were used.

⁵AK Herten, Heidelberg University

6.6. Bleaching of eGFP and Halo-TMR in fixed Huh7.5-gSEP cells

Huh7.5-pBabe-gSEP cells were seeded in two wells of a cleaned LabTek. One well was left unstained while the other well was stained overnight with 10 nM HaloTag TMR ligand. The next day, both wells were washed three times with full growth medium over one hour, fixed with 4 % PFA for 30 min and washed two more times with growth medium. The so prepared cells were imaged at room temperature on a Nikon TiE TIRF microscope.

Photobleaching of TMR in stained Huh7.5-pBabe-gSEP was recorded over 2000 frames at frame rate of 10 Hz and with a 561 nm laser at a power of 8.1 mW as excitation (power of the laser beam as it leaves the objective). The emission light was cleaned up using a 605/70 bandpass filter. Photobleaching of eGFP in stained and unstained Huh7.5-pBabe-gSEP was recorded over 150 frames at frame rate of 20 Hz and with a 488 nm laser at a power of 1.8 mW as excitation (power of the laser beam as it leaves the objective). The emission light was cleaned up using a 525/50 bandpass filter. To introduce a TMR photobleaching step, 50 frames of GFP bleaching were recorded, then TMR was bleached for 2 min with 8.1 mW at 561 nm or eGFP was bleached for 1 min with 4.7 mW at 488 nm. After the bleaching step, another 100 frames were recorded and the two movies (50 frames before the bleaching step and 100 frames after the bleaching step) were concatenated.

For the detection of particles, the Point Source Detection algorithm contained in u-track [36] was slightly modified. The algorithm performs a two-dimensional Gaussian intensity fit to detect potential candidate particles. These candidate particles are further multiplied with a significance matrix to keep only those particles with an intensity that is significantly higher than background. Between these two steps I additionally introduced a dynamic intensity threshold to exclude the dimmest population of candidate particles. For this, I created an intensity histogram of all candidate particles and determined the peak, i.e. the most frequent intensity. At the step of non-significant candidate particle detection, the most abundant particle population represents mostly very dark false-positive detections. Starting from this peak value, I increased the intensity until the frequency of particles with that intensity was smaller than $\frac{1}{30}$ of the peak frequency. This intensity value was defined as intensity threshold.

Using this modified Point Source Detection algorithm, particles were detected in each frame of the recorded movies and the number of particles was plotted over time.

6.7. Determination of the DOL of HaloTag and SNAPf-tag

6.7.1. Beads samples

A Labtek was cleaned as described and multicolor beads (TetraSpeck Microspheres 0.2 μm , Thermo Fisher Scientific) were deposited at different densities per well. The Labtek wells were then filled with ROXS Red, which is an enzymatic oxygen scavenging system combined with reducing and oxidizing agents (ROXS), and sealed with parafilm. ROXS Red was prepared by adding 300 mM Glucose and 12.5 % (v/v) glycerol to 5x PBS buffer and depleting oxygen by vigorously introducing argon into the solution for 20 min with a syringe needle. 1 mM tris(2-carboxyethyl)phosphine, 2 U μM^{-1} glucose oxidase, 250 U μM^{-1} catalase, 1 mM methylviologen and 1 mM ascorbic acid was then added.

6.7.2. Simulation of Single-Molecule Images

To generate images of single emitters with known ground truth, the ISBI Challenge Track Generator provided by the bioimage analysis software Icy [79] was used with the following parameters. Signal-to-noise ratio was set to 8, particle density was set to 500–6600, which translates into particle densities of 0.2–2.6 particles/ μm^2 . Image size was set to 358 x 358 pixel, image depth was 1 and sequence length was 2. For the advanced parameters, xy pixel size was set to 104 to mimic the pixel size of the Nikon TiE TIRF Microscope. All other parameters are irrelevant for immobile particles. Of the resulting two frames, only frame one was saved and ImageJ was used to add additional Gaussian noise with a standard deviation of 10 at an image depth of 8 bit. The ground truth coordinates were saved as XML file and imported into Matlab. A custom code was used to compare ground truth coordinates with the coordinates of detected particles and calculate the relative colocalization as the number of colocalized particles divided by the number of simulated particles. In addition the fraction of false positive particles was determined as the number of false detections (detected but non-colocalized particles) divided by the number of detected particles.

6.7.3. Cell Staining for the Determination of the DOL of HaloTag and SNAPf-tag

Huh7.5 cells expressing the gSEP construct or the LynG construct were seeded into a cleaned LabTek at a density of 20 000–30 000 cells in 200 μL growth medium per well one day before the experiment. In order to test a total of 40 different labeling conditions, every well of the LabTek was stained with a different concentration of

dye substrates (0 nM, 0.1 nM, 1 nM, 5 nM, 10 nM, 50 nM, 100 nM and 250 nM) while five different LabTeks were used for five different incubation times (15 min, 30 min, 1 h, 3 h and 16 h). To measure the overnight labeling conditions, cells were stained shortly after seeding on the day before the experiment. For the other incubation times, cells were grown overnight and stained the next day. Based on a 120 μM stock solution of SiR-BG in DMF and a 80 μM stock solution of HaloTag TMR ligand in DMF, three staining solutions in growth medium were created via serial dilution that contained both dye substrates at a concentration of 1 μM , 100 nM and 10 nM, respectively. These three solutions allowed to set the eight different dye concentrations by adding easily reproducible pipetting volumes between 2 μL and 67 μL of the appropriate staining solution to the 200 μL growth medium already present in each well. After adding the staining solution, medium was pipetted up and down twice to ensure a homogeneous dye concentration within the well and cells were incubated at 37 °C and 5 % CO_2 for the respective incubation time. To remove unbound dye substrates, cells were washed three times with growth medium over a total of 70 min. The washing procedure included two short 15 min washing steps to deplete the medium of unbound dye substrate and a longer 40 min washing step to allow unbound dye to diffuse out of the cell. After the three washing steps, cells were fixed with 4 % PFA for 40 min and washed two times with growth medium. Cells were then ready to be imaged. Although being fixed, cells were imaged in growth medium in order to more closely mimic the situation of live-cell imaging, considering that the photophysics of the dye substrates may well be influenced by the buffer. In fact, both TMR and SiR appeared to show less blinking when imaged in growth medium as compared to imaging in DPBS.

6.7.4. Three-Color Single-Molecule Microscopy of fixed cells

Fixed cells were imaged in growth medium at room temperature. To sustain the largest possible field of view, fixed cells were imaged in full frame mode, i.e. the emission light was projected to the full chip of the camera without being spectrally split. Hence, eGFP, TMR and SiR were sequentially imaged with 1.8 mW at 488 nm, 1.6 mW at 561 nm and 4.8 mW at 640 nm, respectively. eGFP emission and TMR emission were cleaned up using a 525/50 and a 605/70 bandpass filters, respectively, while no bandpass filter was used for SiR emission. For each channel, 100 frames were recorded at 50 ms exposure time and an electronic gain of 100. To avoid pre-bleaching of fluorophores, channels were recorded from long to short wavelength, i.e. from red to blue. After recording the TMR channel, TMR was bleached for 2 min with 8.1 mW at 561 nm in order to exclude Förster resonance energy transfer (FRET) between eGFP and TMR. For each staining condition, 9–13 movies were recorded. After every four conditions, a multicolor bead sample was

imaged in all three channels to serve as a possible registration sample. gSEP and LynG cells were imaged with identical parameters.

6.7.5. Analysis of DOL

Most analysis steps for the determination of the degree of labeling are described in detail in the respective sections of the results. Therefore this section only specifies some basic calculations used during the analysis.

Samples for random colocalization of particles were generated by rotating detected particles in one channel 90° clockwise:

$$\begin{aligned} X_{rotated} &= Y \\ Y_{rotated} &= 512 - X \end{aligned} \quad (6.2)$$

, where X and Y are the x and y coordinates of particles in an image of 512x512 pixel.

To determine the correlation between two populations *A* and *B* (e.g. registration parameters or expression levels of the gSEP construct), Pearson's correlation coefficient was calculated as

$$(R)(A, B) = \frac{1}{N-1} \sum_{i=1}^N \left(\frac{\widehat{A_i} - \mu_A}{\sigma_A} \right) \left(\frac{\widehat{B_i} - \mu_B}{\sigma_B} \right) \quad (6.3)$$

using the MATLAB function *corrcoef*. Correlations were tested for significance using a Student's t distribution for a transformation of the correlation.

Differences between the DOL of HaloTag and SNAPf-tag were tested for significance using a Wilcoxon ranksum test implemented in the MATLAB function *ranksum* with a significance level of $\alpha = 0.05$.

Similarly, differences between the gSEP expression levels determined via Halo-TMR or SNAPf-SiR particles were tested for significance using a Wilcoxon ranksum test implemented in the MATLAB function *ranksum* with a significance of $\alpha = 0.05$.

6.8. Single Molecule Tracking of Labeled Receptors

6.8.1. Cell Staining for Single-Molecule Tracking

Huh7.5 cells expressing Halo-IFNAR1 and SNAPf-IFNAR2 were seeded into a cleaned LabTek at a density of 20 000–30 000 cells in 180 μ L growth medium per well one day before the experiment. As H838 cells generally appeared to grow less

well on glass, H838 cells expressing c-Met-Halo and EGFR-SNAPf were seeded two days before the experiment at a density of 10 000–20 000 cells per well in 160 μL growth medium. Cells were kept at 37 °C and 5 % CO_2 . Wildtype (wt) Huh7.5 or wt H838 cells were used as negative labeling control for every experiment. For H838 cells, 20 μL of a 100 nM solution of SiR-BG in DMEM and 20 μL of a 100 nM solution of HaloTag TMR ligand in DMEM were added to the respective wells around 16 h before the experiment to yield a final concentration of 10 nM of each dye substrate. For Huh7.5 cells, a final concentration of 5 nM SiR-BG and 5 nM HaloTag TMR ligand was obtained by adding 10 μL of a 100 nM solution of each substrate. A homogenous dye concentration was ensured by slowly pipetting up and down the medium two times. Cells were stained overnight at 37 °C and 5 % CO_2 . The next day, the medium was replaced with 250 μL DMEM supplemented with 1 mM sodium pyruvate and 2 mM L-glutamine or Glutamax but without FBS. This medium will be referred to as starvation medium. In addition to remove unbound dye substrate, starvation medium also depletes the cells of growth factors and other FBS-originated signaling molecules that could possibly interfere with the receptors of interest. Cells were incubated in starvation medium at 37 °C and 5 % CO_2 for 15 min. This 15 min washing step with was repeated once more and followed by a 40 min washing step, to allow unbound dye to diffuse out of the cell.

Huh7.5-pBABE-gSEP cells as a control for dimerization were stained overnight with either 10 nM HaloTag TMR ligand and 10 nM SiR-BG (as a positive dimerization control) or 10 nM HaloTag TMR ligand and 10 nM HaloTag SiR-chloroalkane ligand (as a negative dimerization control). Cells were then washed as described above.

6.8.2. Two-Color Live Cell Single-Molecule Tracking

For two-color tracking of membrane receptors, stained live cells were imaged in starvation medium in a heating chamber set to 37 °C. Live Huh7.5-pBABE-gSEP cells as a dimerization control were imaged at room temperature. For all cells, TMR and SiR were simultaneously excited with 1.2 mW at 561 nm and 1.5 mW at 640 nm. The emission light was spectrally split and cleaned up using a 605/70 bandpass filter for TMR and a 685/50 bandpass filter for SiR emission (both AHF Analysetechnik). Each channel was projected to half of the camera chip. The resulting field of view was 256 x 512 pixel with a pixel size of 104 nm. All two-color tracking movies were recorded at a frame rate of 50 Hz (20 ms exposure time), using an electronic gain of 100. The total movie length ranged from 500 to 1000 frames. To allow for a registration of both spectral channels, a multicolor bead (TetraSpeck Microspheres 0.2 μm , Thermo Fisher Scientific) fiducial marker calibration sample was imaged before every experiment.

6.8.3. Bootstrap Significance Test of Dimer Fraction.

To test the difference of the observed dimer fraction between different stimulation conditions for significance, a bootstrap algorithm was applied. First, the t-value for the observed dimer fractions $x_1 \dots x_{n_1}$ and $y_1 \dots y_{n_2}$ of two conditions X and Y was calculated as

$$t_{\text{obs}} = \frac{|\bar{x} - \bar{y}|}{\hat{\sigma}_{\text{pool}} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad (6.4)$$

$$\hat{\sigma}_{\text{pool}} = \sqrt{\frac{(n_1 - 1)\hat{\sigma}_x^2 + (n_2 - 1)\hat{\sigma}_y^2}{n_1 + n_2 - 2}}$$

Under the null hypothesis, X and Y are samples of the same underlying population, hence $\bar{x} = \bar{y}$ and $t = 0$. To create a population that resembles the null hypothesis, I combined all observations from both samples into a common set $\{x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}\}$. From this set, I randomly drew n_1 and n_2 values to form new sets X_b^* and Y_b^* for 10^4 bootstrap cycles. For each two bootstrap sets X_b^* and Y_b^* , I calculated t_b^* following equation 6.4 to create a t-distribution under the null hypothesis. I compared this distribution with t_{obs} and calculated a significance level as the fraction of bootstrap t-values t_b larger or equal t_{obs} .

$$p = \frac{\#\{t_b^* \geq t_{\text{obs}}\}}{10^4} \quad (6.5)$$

6.8.4. Bootstrap Significance Test of Dissociation Rates and Calculation of the Effect Size.

To test the difference between the dissociation rates of different conditions for significance, I applied a combination of two bootstraps. To generate a distribution of dissociation rates for two different conditions, I picked random samples of size n_1 and n_2 from the observed dimer durations $x_1 \dots x_{n_1}$ and $y_1 \dots y_{n_2}$, respectively, and performed a two-component exponential fit as described in section 3.2.3 on page 81 to derive dissociation rates k for each sample. For the HMM-derived dimer durations, I converted the standard histograms to probability densities as described in [82] using equation 4.1 on page 95 and performed a one-component exponential fit on the resulting probability distribution. I repeated this for a total of 100 bootstrap cycles. Based on the resulting distributions $K_1 = k_{1,1} \dots k_{1,100}$ and $K_2 = k_{2,1} \dots k_{2,100}$, I calculated the observed t-value as described above. In addition, I calculated *Cohen's D* in order to evaluate the effect size between the different conditions as

$$D = \frac{|\bar{K}_1 - \bar{K}_2|}{\hat{\sigma}_{pool}} \quad (6.6)$$

$$\hat{\sigma}_{pool} = \sqrt{\frac{(n_1 - 1)\hat{\sigma}_{K_1}^2 + (n_2 - 1)\hat{\sigma}_{K_2}^2}{n_1 + n_2 - 2}}$$

To evaluate the significance of the observed t-value, I combined both distributions to a common set $\{k_{1,1}, \dots, k_{1,100}, k_{2,1}, \dots, k_{2,100}\}$ that resembled the null hypothesis. From this common set I drew random samples K_1^* and K_2^* of size 100 and calculated the t-value t_b^* as described above for a total of 10^3 bootstrap cycles to create a t-distribution under the null hypothesis. I compared this distribution with t_{obs} and, as above, calculated a significance level as the fraction of bootstrap t-values t_b larger or equal t_{obs} .

6.9. Time-lapse Imaging to Determine Receptor Internalization

H838-EGFR-SNAPf/c-Met-Halo cells were seeded into a clean LabTek in full growth medium two days before the experiment. The next day, the cells were stained with 10 nM HaloTag TMR ligand and 10 nM SiR-BG overnight. Cells were then washed three times with starvation medium over 90 min. Starvation medium was then replaced with Leibovitz's L15 medium and cells were imaged at 37 °C at a Nikon TiE TIRF microscopy. All stimulation conditions were imaged in separate LabTeks in order to ensure equal starvation and imaging durations for all conditions. For each condition, 15–20 cells were selected and SNAPf-EGFR and c-Met-Halo were sequentially imaged every minute for 30 min at 1.2 mW 561 nm and 1.5 mW 640 nm in TIRF mode. To induce heterodimerization, I stimulated the cells with either 40 ng/mL EGF, 40 ng/mL HGF or 40 ng/mL of both ligands at the start of the experiment. Before and after each experiment, transmission images were recorded to check cell morphology and determine a ROI for each cell. Within every ROI, the exponential decay of the mean fluorescence intensity yielded the receptor removal rate. Significance was tested using Student's t-test. This analysis was performed by Florian Salopiata.

Chapter 7.

Equipment

7.1. Nikon TIRF microscope

All microscopy data was recorded on an inverted microscope (Nikon TiE) that was equipped with Nikon's PFS2 autofocus system. As light source, a four channel diode/DPSS laser (iChrome MLE-LFA, Toptica) was fiber coupled to Nikon's TIRF-Illuminator, which allows to change the angle under which the laser enters the objective. In this excitation path, a quadruple dichroic mirror (R405/488/561/635, AHF Analysetechnik) separates the laser excitation lines from emission light of the sample. The emission light was collected with a 100x magnification oil objective lens with a numerical aperture of 1.49 and temperature correction (Nikon Apo TIRF 100x 1.49 Oil) in combination with an additional 1.5x magnification lens in the emission path. An quadruple notch filter (ZET405/488/561/640, AHF Analysetechnik) in the emission path blocked remaining laser excitation light. The microscope was further equipped with a spectral image splitter (Optosplit III, Cairn Research) that can create up to three spectral channels. The emission light was finally detected with a 512x512 pixel electron multiplying charge-coupled device (emCCD) camera (Andor iXon+ 897 Ultra, Andor Technology). The resulting pixel size was 104 nm. For single-molecule tracking of live cells, the microscope was equipped with a custom-built heating chamber and a objective heating ring.

7.2. Software

All MATLAB analysis was performed with MATLAB R2015a. For automated image editing, I used the ImageJ based software Fiji [78]. Particle detection and tracking was performed based on u-track [36]. Geneious 8.1 (Biomatters) was used to design PCR primers and generally plan cloning steps. Figures were finalized with the use of Inkscape 0.92.1.

Part V.
Appendix

Appendix A.

Appendix

A.1. Cloning of the gSEP and LynG construct

SNAP26m was bought from New England Biolabs, hence its DNA sequences was available for synthesis. To transfer the gSEP precursor from pIDTSmart into pMOWS, I digested both vectors with EcoRI and PacI, purified the pMOWS vector backbone and the gSEP precursor insert via agarose gel extraction before ligating them. I transformed competent e.coli of the DH5- α strain with the resulting plasmid (pMOWS-SEP0) and grew single colonies on LB-agar containing $100 \mu\text{g mL}^{-1}$ ampicillin. Upon miniprep and test-digestion with AgeI and NotI, all tested colonies gave the expected bands of 5264 bp and 1281 bp (see Figure A.1 (a)). pMOWS-SEP0 was then amplified by Maxiprep and finally confirmed by sequencing.

To inserted HaloTag between the BamHI and AgeI restriction sites, I amplified the DNA sequence of HaloTag by PCR from a pMOWS-Tyk2-Halo plasmid provided by Ursula Klingmüller using PCR primers with a respective BamHI and AgeI overhang^{1,2}. HaloTag was initially bought from Promega within a cooperation project of the Klingmüller and Herten labs. The purified PCR product and the pMOWS-SEP0 plasmid were digested with BamHI and AgeI, purified, ligated and the resulting pMOWS-SEP1 plasmid was introduced into competent DH5- α e.coli. Again, transformed e.coli were grown on ampicillin-containing LB-agar and colonies were picked for miniprep and test-digestion with AgeI and BamHI. All colonies gave the expected insert size of 888 bp (see Figure A.1 (b)). One colony was chosen for maxiprep and the purified pMOWS-SEP1 plasmid was verified by sequencing.

SNAPf-tag was bought from New England Biosciences and amplified with NdeI and MfeI restriction site overhangs via PCR^{3,4}. To replace SNAP26m with SNAPf, the PCR product and the pMOWS-SEP1 plasmid were digested with NdeI and

¹fwd: 5'-CGCGGATCCGCAGAAATCGGTACTGGCTTTCCATTC-3'

²rev: 5'-CGCACCGGTGCCGAAATCTCCAGCGTCGAC-3'

³fwd: 5'-GCGCATATGGACAAAGACTGCGAAATGAAGCGCACC-3'

⁴rev: 5'-GCGCAATTGACCCAGCCAGGCTTGCCAG-3'

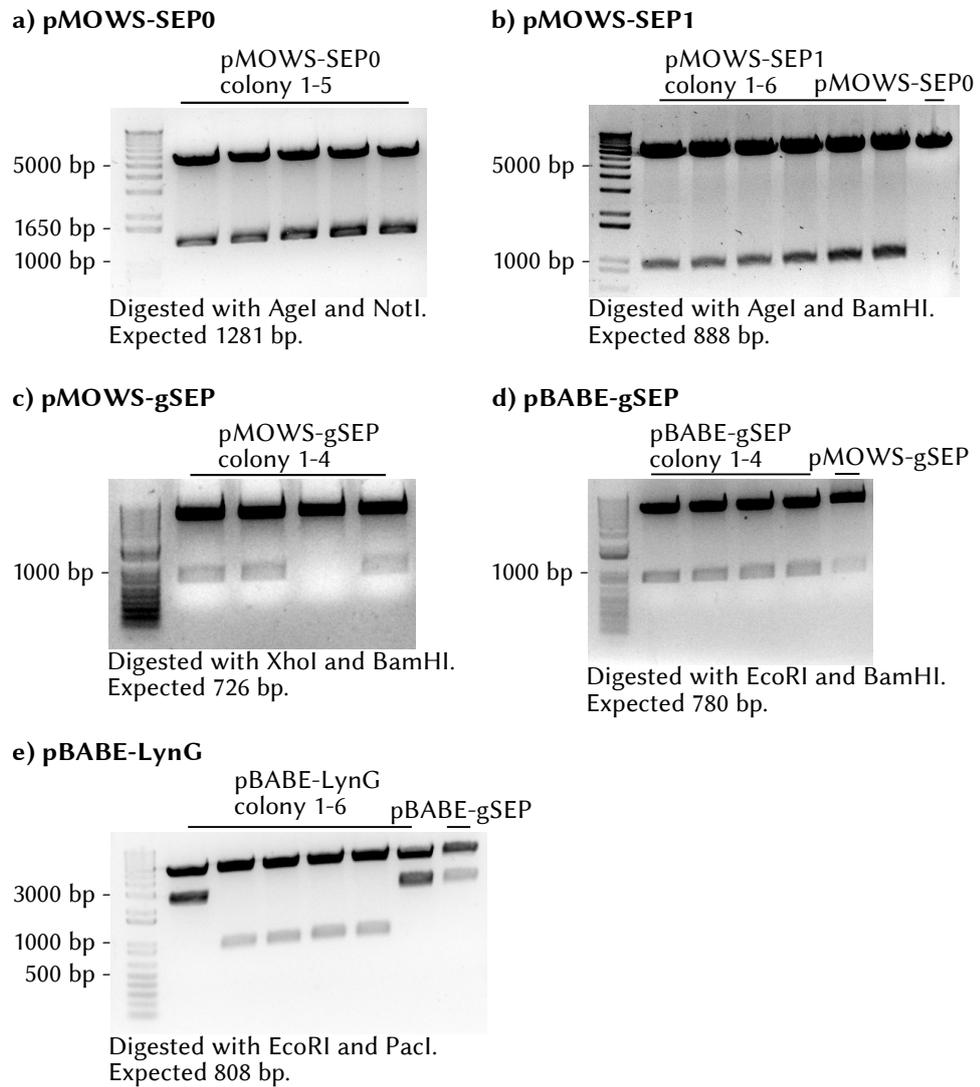


Figure A.1.: Validation of insert sizes for all cloning steps. pBABE-gSEP and pBABE-LynG were stepwise assembled. At the end of each cloning step, the plasmids from transformed e.coli DH5 α colonies were checked for the correct insert size. A colony with the correct plasmid was chosen for maxiprep and the final purified plasmid was verified by sequencing

MfeI, purified and ligated. Transformed *e.coli* DH5- α were grown on ampicillin-containing LB-agar and four colonies were picked for miniprep. Since SNAP26m-tag and SNAPf-tag both have a size of 546 bp, the correct replacement cannot be judged based on a test-digestion. Therefore the purified plasmids from all four colonies were sequenced and SNAPf was detected in one of the four colonies. This colony was used for maxiprep and the resulting pMOWS-SEP1f plasmid was again confirmed by sequencing.

To insert eGFP between the XhoI and BamHI restriction sites of pMOWS-SEP1f, I amplified the eGFP DNA sequence via PCR with XhoI and BamHI overhangs using a HaloTag-eGFP-Tubulin plasmid provided by Klaus Yserentant⁵ as a template^{6,7}. Digestion of both the PCR product and the pMOWS-SEP1f plasmid with XhoI and BamHI followed by purification and ligation produced the pMOWS-gSEP plasmid (see Figure 2.2 (c)). *E.coli* DH5- α were transformed and grown on ampicillin-containing LB-agar. Three out of four picked colonies showed the expected insert size of 726 bp after miniprep and test-digestion with XhoI and BamHI (see Figure A.1 (c)). pMOWS-gSEP from these three colonies was pooled and used for further cloning.

To transfer the gSEP construct into the lower expressing retroviral pBABE plasmid, I extracted gSEP from pMOWS-gSEP via digestion with EcoRI and PacI and inserted it in the pBABE plasmid using the same restriction sites. A plasmid of the correct size could be detected by test-digestion with EcoRI and BamHI in all transformed *e.coli* DH5- α colonies (see Figure A.1 (d)). Again, due to their similar size, pBABE-gSEP and pMOWS-gSEP can not be differentiated on size alone. Therefore, one colony was picked for maxiprep and pBABE-gSEP was validated against pMOWS-gSEP by sequencing.

pBABE-LynG was created from pBABE-gSEP by restriction digestion with BamHI and MfeI, blunting the sticky ends with T4 DNA Polymerase and ligating the plasmid again. Because there is a low probability that the HaloTag and SNAPf-tag insert re-ligates with the plasmid despite blunting, six of the transformed *e.coli* DH5- α colonies were picked, their respective plasmids were purified via miniprep and a test digestion with EcoRI and PacI was performed. Four out of six colonies gave the expected insert size of 808 bp while two colonies contained the undigested or re-ligated gSEP insert of 2328 bp (see Figure A.1 (e)). To obtain pBABE-LynG, a maxiprep of colony 2 was performed and the purified pBABE-LynG plasmid was validated by sequencing.

⁵AK Herten, Heidelberg University

⁶fwd: 5'-ATCCCTCGAGACCATGGTGAGCAAG-3'

⁷rev: 5'-CACGGATCCCTTGTACAGCTCGTC-3'

A.2. Dimerization-dependent reduction of particle mobility

To analyze if dimerization had an effect on particle diffusion, I determined the jump size of candidate particles for all time steps of a candidate trajectory pair for all conditions. Figure A.2 (a) shows the pooled jump size distribution of IFNAR1 and IFNAR2 receptors upon stimulation with 1 nM IFN α 2. Figure A.2 (c) shows the pooled jump size distribution of EGFR and c-Met receptors upon stimulation with 40 ng/mL EGF and HGF. For both receptor systems, the distribution of dimer particles is clearly shifted to shorter jump sizes when compared to free particles. For all stimulation conditions, the median jump size is reduced by approximately 20 % (Fig A.2 (b) and (d)). This reduction of the median jump size from free to dimer particles was highly significant for all conditions (Fig A.2 (b)).

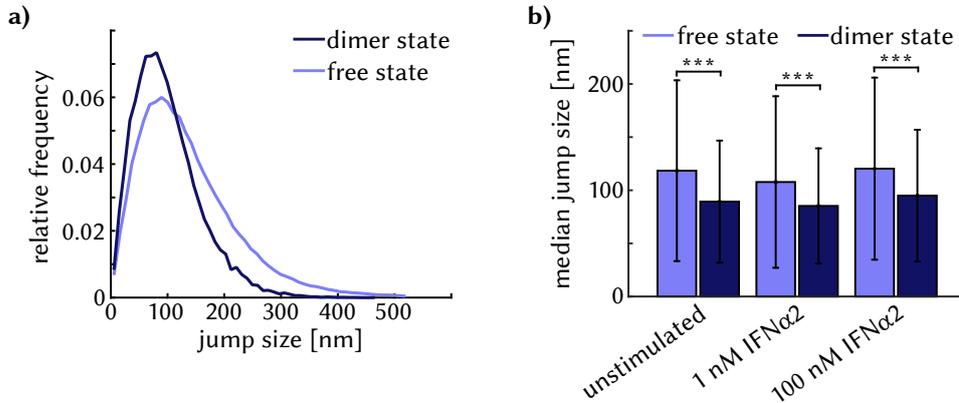
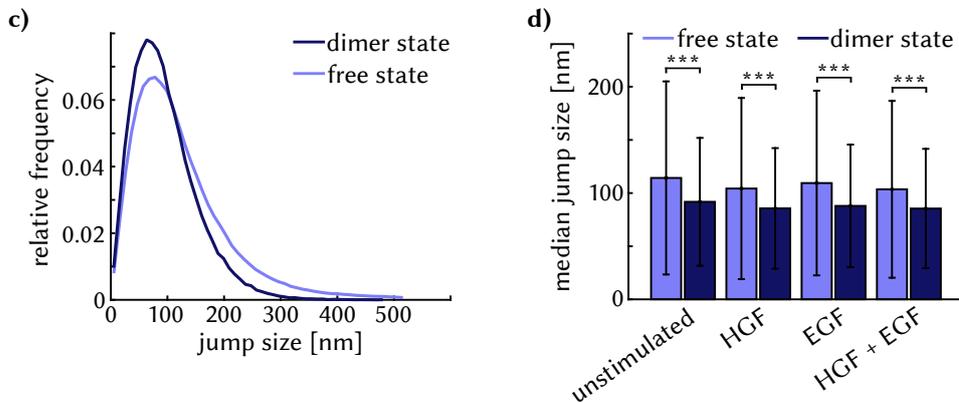
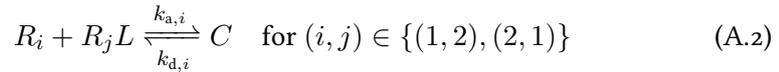
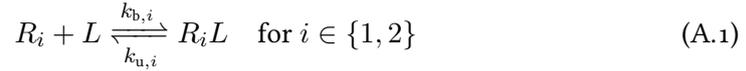
IFNAR1/IFNAR2 heterodimerization**EGFR/c-Met heterodimerization**

Figure A.2.: Jump sizes of free and dimer particles. **a),c)** Jump size distributions of candidate particles, (a) IFNAR1 and IFNAR2 combined from Huh7.5-Halo-IFNAR1/SNAPf-IFNAR2 cells stimulated with 1 nM IFN α 2, (b) EGFR and c-Met combined from H838-EGFR-SNAPf/c-Met-Halo cell stimulated with 40 ng/mL EGF and HGF. The jump sizes of particles classified as 'dimer' are shifted toward short jump sizes when compared to 'free' particles. **b),d)** For all conditions, the median jump size for dimer particles was significantly lower than the median jump size of free particles. The error bars represent the standard deviation. The number of analyzed IFNAR jump sizes was between 92 096–280 204 for free particles and between 22 538–52 240 for dimer particles. The number of analyzed EGFR and c-Met jump sizes was between 408 728–1 263 264 for free particles and between 50 598–145 424 for dimer particles. The significance level was $p < 0.001$ (***) based on bootstrap significance testing (analogous to section 6.8.3 on page 123 of material and methods).

A.3. Bell-shaped equilibrium binding curves for dimerizing receptor systems

Nikolas Schnellbacher⁸ has set up a kinetic model for the ligand-dependent formation of the Interferon- α receptor (IFNAR) ternary complex formation. It is based on the following reaction



, where R_i stands for the receptor subunit i , L denotes the ligand, $R_i L$ represents the (binary) bound state and $C = R_1 L R_2 = R_2 L R_1$ denoted the ternary complex. The rate coefficients $k_{b,i}$ and $k_{u,i}$ describe the binding and unbinding kinetics of the ligand from a receptor subunit i , while $k_{a,i}$ and $k_{d,i}$ describe the assembly and disassembly rate of the ternary complex C from/into a binary complex $R_j L$ and an individual receptor chain R_i ($i \neq j$). It should be noted that equation (1) describes a 3D kinetic between ligands in bulk solution and membrane-associated surface receptor. in contrast, equation (2) describes the 2D reaction between receptors that is restricted to the plasma membrane. In chemical equilibrium the equilibrium constants can be obtained as

$$K_{D,i}^{3d} = \frac{k_{u,i}}{k_{b,i}}, \quad K_{D,i}^{2d} = \frac{k_{d,i}}{k_{a,i}}, \quad \text{for } i \in \{1, 2\}. \quad (\text{A.3})$$

To describe the receptor assembly in the homogeneous well-mixed limit, the reaction schemes (1) & (2) were converted into a set of kinetic rate equations governed by mass action kinetics. These kinetic rate equations were analytically solved in the steady state and the fraction of ternary complexes f was expressed as a function of the receptor density R_0 , the ligand concentration L and the equilibrium dissociation constants $\{K_{D,i}\}$. For this, the steady state equations were rewritten as a system of nonlinear equations as a function of the equilibrium dissociation constants $\{K_{D,i}\}$, the ligand concentration $[L]$ and the receptor density R_0 and isolate for the steady state solution of the ternary complexes $[C]_s$.

Based on the equilibrium dissociation constants

$$K_{D,i}^{3d} = \frac{k_{u,i}}{k_{b,i}} = \frac{[L][R_i]}{[R_i L]} \quad \text{for } i \in \{1, 2\}. \quad (\text{A.4})$$

⁸AG Schwarz, Heidelberg University

for the $3d$ reactions and

$$K_{D,i}^{2d} = \frac{k_{d,i}}{k_{a,i}} = \frac{[R_i][R_jL]}{[C]} \quad \text{for } (i, j) \in \{(1, 2), (2, 1)\}. \quad (\text{A.5})$$

for the $2d$ interactions, the result for $[C]_s$ were used to calculate the solution for the fraction of ternary receptors f :

$$\begin{aligned} f := \frac{[C]_s}{R_0} = f([L], R_0, \{K_{D,i}\}) &= \frac{1}{2[L]K_{D,1}^{3d}} \left[K_{D,1}^{3d}K_{D,2}^{3d}K_{D,1}^{2d} + \right. \\ &\quad \left. (K_{D,2}^{3d} + K_{D,1}^{3d} + [L]) K_{D,1}^{2d}[L] + 2K_{D,1}^{3d}[L]R_0 - \right. \\ &\quad \left. \sqrt{K_{D,1}^{2d}(K_{D,1}^{3d} + [L])(K_{D,2}^{3d} + [L])} \times \right. \\ &\quad \left. \times \sqrt{K_{D,1}^{3d}K_{D,2}^{3d}K_{D,1}^{2d} + (K_{D,2}^{3d} + K_{D,1}^{3d} + [L]) K_{D,1}^{2d}[L] + 4K_{D,1}^{3d}[L]R_0} \right] \quad (\text{A.6}) \end{aligned}$$

The fraction of ternaries f further allows to iteratively solve for the fraction of intermediate species:

$$f_{R_1L} = \frac{[R_1L]_s}{R_0} = (1 - f) \cdot \frac{[L]}{K_{D,1}^{3d} + [L]} \quad (\text{A.7})$$

$$f_{R_2L} = \frac{[R_2L]_s}{R_0} = \left(\frac{K_{D,1}^{2d}}{R_0} \right) \frac{f}{1 - f_{R_1L} - f} \quad (\text{A.8})$$

$$f_{R_1} = \frac{[R_1]_s}{R_0} = 1 - f_{R_1L} - f \quad (\text{A.9})$$

$$f_{R_2} = \frac{[R_2]_s}{R_0} = 1 - f_{R_2L} - f \quad (\text{A.10})$$

The result for f viewed as a function of the stimulation $[L]$ is typically referred to as bell-shaped equilibrium binding curve, when plotted in logarithmic concentration space (Fig A.3).

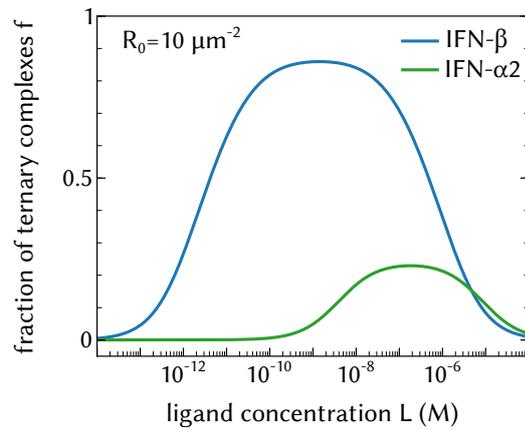
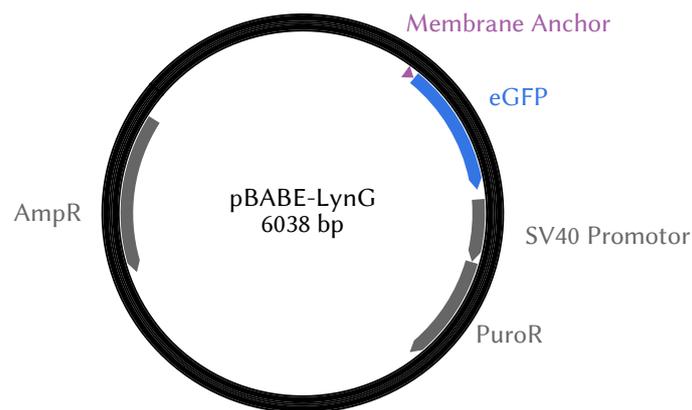


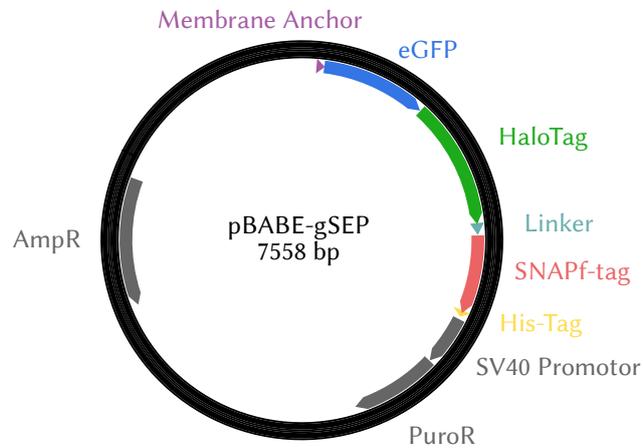
Figure A.3.: Equilibrium binding curve of ternary complexes The fraction of ternary complexes as a function of ligand concentration and receptor density describes a bell-shaped equilibrium binding curve. While low to moderate ligand concentrations induce ternary complex formation, ligand oversaturation reduces the fraction of ternary complexes. The sensitive ligand concentration range depends on the receptor density as well as on the reaction rates between the complex components.

A.4. Plasmid Sequences

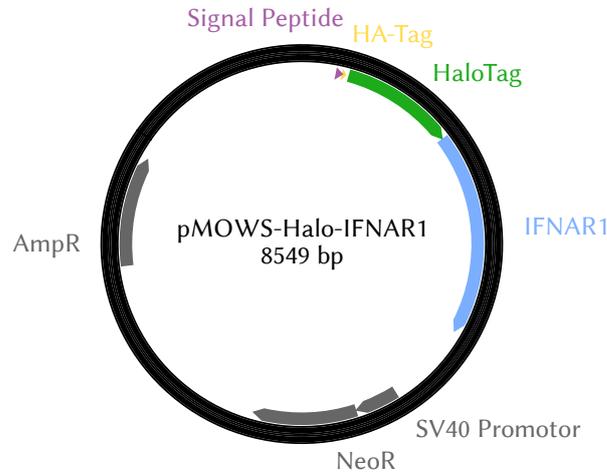


```

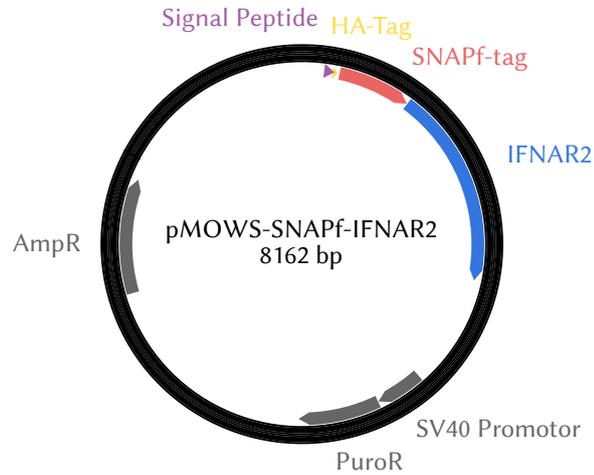
ATGGGATGTATCAAGAGTAAGCGTAAGGATAATCTCAATGACGACGAGCTCGAGACCATGGTGAGCAAG
GGCGAGGAGCTGTTCAACGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACA
AGTTCAGCGTGTCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTG
CACCACCGGCAAGCTGCCCGTGGCCCTGGCCACCCTCGTGACCACCCTGACCTACGGCGTGCGAGTGC
TTCAGCCGCTACCCCGACCATGAAGCAGCAGCACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGT
CCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAG
GGCGACACCCTGGTGAACCGCATCGAGCTGAAGGCATCGACTTCAAGGAGGACGGCAACATCCTGG
GGCACAAGCTGGAGTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCA
TCAAGGTGAAGTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAG
CAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCCGACAACCACTACCTGAGCACCCAGTCCA
AACTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCGG
GATCACTCTCGGCATGGACGAGCTGTACAAG
  
```



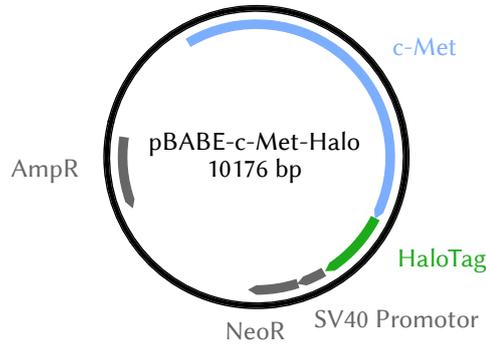
ATGGGATGTATCAAGAGTAAGCGTAAGGATAATCTCAATGACGACGAGCTCGAGACCATGGTGAGCAAG
GGCGAGGAGCTGTTACCGGGGTGGTCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACA
AGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACTACGGCAAGCTGACCCTGAAGTTCATCTG
CACCACCGGCAAGCTGCCCGTGCCTGGCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGC
TTCAGCCGCTACCCCGACCACATGAAGCAGCAGACTTCTCAAGTCCGCCATGCCCGAAGGCTACGT
CCAGGAGCGCACCATTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAG
GGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGG
GGCACAAGCTGGAGTACAAC TACAACAGCCACAACGTCTATATCATGGCCGACAAGCAAGAACGGCA
TCAAGGTGAAGTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAG
CAGAACACCCCATCGGCGACGGCCCGTGTGTGTCGCCGACAACCACTACCTGAGCACCAGTCCA
AACTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCTCTGCTGGAGTTCGTGACCGCCGCGG
GATCACTCTCGGCATGGACGAGCTGTACAAGGGATCCGCAGAAATCGGTACTGGCTTTCATTGACCC
CCATTATGTGGAAGTCTGGGCGAGCGCATGCACTACGTCGATGTTGGTCCGCGCGATGGCACCCTG
TGCTGTTCTGCACGGTAACCCGACCTCCTCCTACGCTGTGGCGCAACATCATCCCGCATGTTGCACCGA
CCCATCGCTGCATTGCTCCAGACCTGATCGGTATGGGCAAATCCGACAACCCAGACCTGGTTATTTCT
TCGACGACCACGTCCGCTTCATGGATGCCCTTCATCGAAGCCCTGGGTCTGGAAGAGGTGCTCCTGGT
ATTCACGACTGGGGCTCCGCTCTGGGTTTCCACTGGGCAAGCGCAATCCAGAGCGCGTCAAAGGTAT
TGCATTTATGGAGTTCATCCGCCCTATCCCGACCTGGGACGAATGGCCAGAATTTGCCCGGAGACCTT
CCAGGCCCTCCGCACCACCGACGTCGGCCGCAAGCTGATCATCGATCAGAACGTTTTTATCGAGGGTA
CGCTGCCGATGGGTGTCGTCCGCCGCTGACTGAAGTCGAGATGGACCATTACCGGAGCCGTTTCT
GAATCCTGTTGACCGGAGCCACTGTGGCGCTTCCCAAACGAGCTGCCAATCGCCGGTGAGCCAGCG
AACATCGTCGCGCTGGTCAAGAATACATGGACTGGCTGCACCAGTCCCTGTCCCGAAGCTGCTGTT
CTGGGGACCCAGGCGTCTGATCCACCCGGCGAAGCCGCTCGCTGGCCAAAAGCCTGCCTAAC
TGCAAGGCTGTGGACATCGGCCCGGGTCTGAATCTGCTGCAAGAAGACAACCCGGACCTGATCGGCA
GCGAGATCGCGCGCTGCTGTCGACGCTGGAGATTTCCGGCACCGGTTTGGCGGAGGCGGGCGGA
AGGAGGCGCGGCGGAAGGAGGCGCGGCGGAAGGAGGCGCGGCGGAAGGCGGCGCGCATATGGAC
AAAGACTGCGAAATGAAGCGCACACCCTGGATAGCCCTCTGGGCAAGCTGGAAGTGTCTGGGTGCGA
ACAGGGCCTGCACCGTATCATCTTCTGGGCAAAGGAACATCTGCCCGGACGCGCTGGAAGTGCCTG
CCCCAGCCCGCTGCTGGGCGGACAGAGCCACTGATGCAAGGCCACCGCCTGGCTAACGCCACTT
TCACCAGCCTGAGGCCATCGAGGAGTTCCTGTGCCAGCCCTGCACCACCCAGTGTCCAGCAGGAG
AGCTTTACCCGCCAGGTGCTGTGAAACTGCTGAAAGTGGTGAAGTTCGGAGAGGTCATCAGCTACAG
CCACCTGGCCGCCCTGGCCGCAATCCCGCCGCCACCGCCGCGTGAACCCGCTGAGCGGAAA
TCCCGTGCCATTCTGATCCCCTGCCACCGGGTGGTGCAGGGCGACCTGGACGTGGGGGGTACGAG
GGCGGGCTCGCCGTAAAGAGTGGCTGCTGGCCACGAGGGCCACAGACTGGGCAAGCCTGGGCTG
GGTCAATTGCACCATCACCATCACCAC



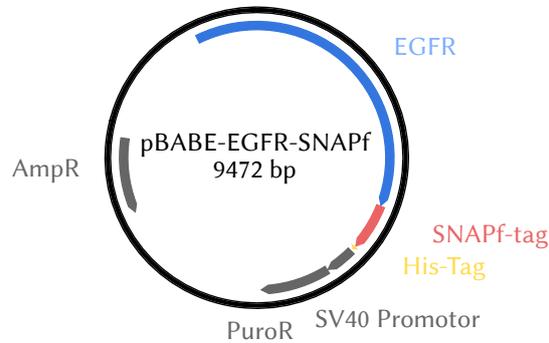
ATGGAGACAGACACTCTGCTATGGGTAAGTCTGCTGCTGGGTTCCAGGTTCCACTGGTGACTATCCAT
 ATGATGTTCCAGATTATGCTGGGGCCAGCCGGCCAGATCTGGATCCGAAATCGGTACTGGCTTTCCAT
 TCGACCCCATATATGTGGAAGTCTGGGGCAGCGCATGCACTACGTCGATGTTGGTCCGCGCGATGGC
 ACCCTGTGCTGTTCCCTGCACGGTAACCCGACCTCCTCCTACGTGTGGCGCAACATCATCCCGCATGTT
 GCACCGACCCATCGCTGCATTGCTCCAGACCTGATCGGTATGGGCAAATCCGACAAACCAGACCTGGG
 TTATTTCTTCGACGACCACGCTCCGCTTCATGGATGCCTTCATCGAAGCCCTGGGTCTGGAAGAGGTCGT
 CCTGGTCATTACGACTGGGGCTCCGCTCTGGGTTTCCACTGGGCAAGCGCAATCCAGAGCGCGTC
 AAAGGTATTGCATTATGGAGTTTCATCCGCCCTATCCGACCTGGGACGAATGGCCAGAATTTGCCCGC
 GAGACCTTCCAGGCCTTCCGCACCACCGACGTCGGCCGCAAGCTGATCATCGATCAGAACGTTTTTAT
 CGAGGTACGCTGCCGATGGGTGTCGTCGCCCGCTGACTGAAGTCGAGATGGACCATTACCGCGAG
 CCCTTCTGAATCCTGTTGACCGCGAGCCACTGTGGCGCTTCCCAAACGAGCTGCCAATCGCCGGTGA
 GCCAGCGAATCGTCGCGCTGGTGAAGAATACATGGACTGGCTGCACCAGTCCCTGTCCCGAAGC
 TGCTGTTCTGGGGCACCCAGGCGTTCTGATCCACCGCCGGAAGCCGCTGCGCTGGCCAAAAGCCT
 GCCTAACTGCAAGGCTGTGGACATCGGCCCGGCTGTAATCTGCTGCAAGAAGACAACCCGGACCTGA
 TCGGCAGCGAGATCGCGCGCTGGCTGTCTACTCTGGAGATTTCCGGTAGATCTAAAAATCTAAAATCTC
 CTCAAAAAGTAGAGGTCGACATCATAGATGACAACCTTATCCTGAGGTGGAACAGGAGCGATGAGTCTGT
 CGGGAATGTGACTTTTTTCATTCGATTATCAAAAACTGGGATGGATAATTGGATAAAATGTCTGGGTGC
 AGAATATTACTAGTACCAAAATGCAACTTTTCTTCACTCAAGCTGAATGTTATGAAGAAATAAATTCGCTA
 TAAGAGCAGAAAAAGAAAACACTTCTTCATGGTATGAGGTTGACTCATTACACCATTTCGCAAAGCTCA
 GATTGGTCTCCAGAAGTACATTTAGAAGCTGAAGATAAGGCAATAGTGATACACATCTCTCTGGAAACA
 AAAGATAGTGTATGTGGGCTTTGGATGGTTAAGCTTTACATATAGCTTACTTATCTGGAAAACTCTTCA
 GGTGTAGAAGAAAGGATTGAAAATATTTCCAGACATAAAAATTTATAAACTCTCACCAGAGACTACTTAT
 TGTCTAAAAGTTAAAGCAGCACTACTTACGTCATGGAAAATTTGGTGTCTATAGTCCAGTACATTGTATAAA
 GACCACAGTTGAAAATGAACTACCTCCACCAGAAAATATAGAAGTCAGTGTCCAAAATCAGAACTATGTT
 CTAAATGGGATTATACATATGCAAAATGACCTTCAAGTTCAGTGGCTCCACGCCTTTTTAAAAAGGAA
 TCCTGAAAACCATTTGTATAAATGGAACAAATACCTGACTGTGAAAATGTCAAACTACCCAGTGTGCT
 TTCTCAAAACGTTTTTCAAAAAGGAATTTACCTTCTCCGCGTACAAGCATCTGATGGAAATAACACATC
 TTTTGGTCTGAAGAGATAAAGTTTGATACTGAAATACAAGCTTTCCTACTTCCTCCAGTCTTTAACATTA
 GATCCCTTAGTGATTCAATCCATATCTATATCGGTGCTCCAAAACAGTCTGGAAAACAGCCTGTGATCCA
 GGATTATCCACTGATTATGAAATATTTTTGGGAAAACACTTCAAATGCTGAGAGAAAAATATCGAGA
 AAAAACTGATGTTACAGTTCCTAATTTGAAACCACTGACTGTATATTGTGTGAAAGCCAGAGCACACAC
 CATGGATGAAAAGCTGAATAAAAGCAGTGTTTTTAGTACGCTGTATGTGAGAAAAACAAAACAGGAAAT
 ACCTCTAAAATTTGGCTTATAGTTGGAATTTGTATGCAATATTGCTCTCCCGTTTGTCAATTTATGCTGG
 AAAGTCTTCTTGAGATGCATCAATATGTCTTCTTCCATCACTTAAACCTTCTCCAGTATAGATGAGTAT
 TTCTCTGAACAGCCATTGAAGAATCTTCTGCTTCAACTTCTGAGGAACAATCGAAAAATGTTTCATAAT
 TGAATATAAGCACAATGCTACAGTAGAAGAACTAATCAAATGATGAAGATCATAAAAAATACAGTTCC
 CAAAATAGCCAAGATTTCAGGAAATTTCTAATGAAGATGAAAAGCGAAAAGTAAAACAAGTGAAGAACTA
 CAGCAGGACTTTGTA



ATGGAGACAGACACACTCCTGCTATGGTACTGCTGCTCTGGGTTCCAGGTTCCACTGGTGACTATCCAT
 ATGATGTTCCAGATTATGCTGGGGCCAGCCGGCCAGATCTATGGACAAAGACTGCGAAATGAAGCGCA
 CCACCCTGGATAGCCCTCTGGGCAAGCTGGAAGTGTCTGGGTGCGAACAGGGCTGCACCGTATCATC
 TTCCTGGGCAAGGAACATCTGCCGCCGACGCCGTGGAAGTGCCTGCCCCAGCCGCCGTGCTGGGC
 GGACCAGAGCCACTGATGCAGGCCACCGCCTGGCTCAACGCCTACTTTCACCAGCCCTGAGGGCCATCG
 AGGAGTTCCTGTGCCAGCCCTGCACCACCCAGTGTTCACGAGGAGAGCTTTACCCGCCAGGTGCT
 GTGGAACTGCTGAAAGTGGTGAAGTTCGGAGAGGTCATCAGCTACAGCCACTGGCCGCCCTGGCC
 GGCAATCCCGCCGCCACCGCCGCCGTGAAAACCGCCCTGAGCGGAAATCCCGTGCCATTCTGATCC
 CTGCCACCGGGTGGTGCAGGGCGACCTGGACGTGGGGGGCTACGAGGGCGGGCTCGCCGTGAAAG
 AGTGGCTGCTGGCCCACGAGGGCCACAGACTGGGCAAGCCTGGGCTGGGTAGATCTTCATATGATTCC
 CCTGATTACACAGATGAATCTTGCACTTCAAGATATCATTGCGAAATTTCCGGTCCATCTTATCATGGGA
 ATTAAAAACCCTCCATTGTACCAACTCACTATACATTGCTGTATACAATCATGAGTAAACCAGAAGATTT
 GAAGTGGTTAAGAAGTGTGCAAATACCACAAGATCATTTTGTGACCTCACAGATGAGTGGAGAAGCAC
 ACACGAGGCCTATGTCACCGTCTTAGAAGGATTCAGCGGGAACACAACGTTGTTTTCAGTTGCTCACACAA
 TTTCTGGCTGGCCATAGACATGTCTTTTGAACCACCAGAGTTTGAGATTGTTGGTTTTACCAACCACATT
 AATGTGATGGTGAATTTCCATCTATTGTTGAGGAAGAATTACAGTTTGATTATCTCTCGTCAATTGAAGA
 ACAGTCAGAGGGAATTGTTAAGAAGCATAAACCCGAAATAAAAGGAAACATGAGTGGAAATTTACCTAT
 ATCATTGACAAGTTAATCCAAACACGAAGTGTGTATCTGTTATTTAGAGCACAGTGTGAGCAAGC
 AGTAATAAAGTCTCCCTAAAATGCACCTCCTTCCACCTGGCCAGGAATCAGAATCAGCAGAATCTGCC
 AAAATAGGAGGAATAATTACTGTGTTTTGATAGCATTGGTCTTGACAAGCACCATAGTGACACTGAAATG
 GATTGGTTATATATGCTTAAGAAATAGCCTCCCCAAAGTCTTGAATTTTCATAACTTTTTAGCCTGGCCAT
 TTCCTAACCTGCCACCGTTGGAAGCCATGGATATGGTGGAGGTCATTTACATCAACAGAAAGAAGAAAG
 TGTGGGATATAATTATGATGATGAAAGTGATAGCGATACTGAGGCAGCGCCAGGACAAGTGGCGGTG
 GCTATACCATGCATGGACTGACTGTCAGGCCCTCTGGGTCAGGCCTCTGCCACCTCTACAGAATCCCAGT
 TGATAGACCCGGAGTCCGAGGAGGAGCCTGACCTGCCTGAGGTTGATGTGGAGCTCCCCACGATGCCA
 AAGGACAGCCCTCAGCAGTTGGAACCTTTGAGTGGGCCCTGTGAGAGGAGAAAGAGTCCACTCCAGGA
 CCCTTTCCCGAAGAGGACTACAGCTCCACGAGGGGTCTGGGGCAGAAATTACCTCAATGTGGACT
 TAAACTCTGTGTTTTGAGAGTCTTGATGACGAGGACAGTACGACTTAGAAGCCCTCTGATGCTATC
 GTCTCATCTGGAAGAGATGGTTGACCCAGAGGATCCTGATAATGTGCAATCAACCATTGCTGGCCAG
 CGGGGAAGGGACACAGCCAACCTTTCCAGCCCTCTTCAGAGGGCCTGTGGTCCGAGATGCTCCA
 TCTGATCAAAGTGACACTTCTGAGTCAGATGTTGACCTTGGGGATGGTTATATAATGAGA



ATGAAGGCTCCTGCCGTGCTGGCCCTGGCATCCTGGTGTGCTTACCCTGGTGCAGCGGAGCAACGGCGAGTGCAAAAGAC
CCTGGCCAAGAGCGAGATGAACGTGAACATGAAGTACCAGCTGCCCACTTCACCGCCGAGACACCCATCCAGAACGTGATCCTGCA
CGAGCACCACATCTTTCTGGGGCCACCAACTACATCTACGTGCTGAACGAAGAGGATCTGCAGAAGGTGGCCGAGTACAAGACCGG
CCCTGTGTGGAACACCCCGACTGCTTCCATGCCAGGACTGCAGCAGCAAGGCCAACCTGTCTGGCGCGGTGGGAAGGACAACT
CAACATGGCCCTGGTGGTGGACACTACTACGACGACCACTGATCTCTGGCCAGCGTGAACCGGGCAGCTGCAGAGACAGCT
GTCCCCCACAAACACCCGGCCAGATCCAGAGCGAGGTGCAGTGCATCTTCAAGTCCAGATCGAGGAACCCAGCCAGTGGCCCGA
CTGGCTGGTGTCTCCCTGGGGCCAAAGTGTCTGAGCAGCTGAAGGACAGATTCATCAATTTCTTGGGGCAACCCATCAACAG
CAGTACTTCCCGGACACCCCTGCACGCATCAGCGTGGCGCGGTGAAGAGACAAAGGACGGCTCATGTCTCTGACCGACCA
GAGTACATCGATGTCTGCCGAGTTCGGGACAGTACCCATCAATACGTGCACGCTTCGAGAGCAACAACCTCATCTACTTTC
TGACCGTGCAGAGAGACACTGGACGCCAGACCTTCCACACCCGGATCCCGGTTCTGCAGCATCAACAGCGGCCTGCACAGCT
ACATGGAAATGCCCTGGAATGCATCCTGACCAGAGAGCGGAAGAAGCGGTCCACCAAGAAAGAGGTGTTCAACATCCTGCAGGCCG
CCTACGTGTCCAAGCTGGCCGCCAGCTGGCCAGAGATCGCCGCGCAGCTGAACGACGACATCCTGTTCCGGCTGTTCCGCCAG
AGCAAGCCCGACAGCGCCGAGCCATGGACAGATCCGCCATGTGGCCCTTCCCTATCAATATGTGAACGACTTCTTCAACAAAGATCG
TGAACAAGAACAACGTGGCGTGCCTGCAGCACTTCTACGGCCCAACACGAGCACTGCTTCAACCGGACCCCTGCTGGGAACAGCA
GCGGCTGGCAGGCGAGCGGGACGAGTACCGGACCGAGTTCACACCGCCCTGCAGCGGGTGGACCTGTTATGGGCGAGTTCAG
CGAGGTGCTGTGACCAGCATCAGCACCTTCAAGGGCGACTGACAATCGCCAACCTGGGACCCAGCGAGGGCCGGTTCATGCA
GGTCTGGTGTCCAGAAGCGGCCAGCACCACCCACGTGAACCTTCTGCTGGACAGCCCGTGTCCCGGAAGTGTCTGGT
AACACACCCTGAACAGAACGGCTACACCTGGTTCATACCGGCAAGAAGATCACAAGATCCCTGAAACGGCTGGGCTGGCCG
ACTTCCAGAGCTGCAGCCAGTGTCTGAGCGCCACCCCTTGTGTCAGTGGGCTGGTGCACGATAAGTGGCTGGGAGCGAGGAA
TGCCTGAGCGGCACCTGGACCCAGCAGATCTGCCTGCCCGCATCTACAAGGTGTTCCCAACAGCGCCCTCTGGAAGCGGCAC
CAGACTGACCATCTGCGCTGGGACTTCCGCTTCCGGCGGAACAACAGTTCGACCTGAAGAAAACCCGGGTGCTGTGGGCAAG
AGAGCTGCACCCCTGACCTGAGCGAGAGCACTGAACACCCCTGAAGTGCACCGTGGGACCGCCATGAACAAGCACTTCAACATGA
GCCATCATCAGCAACCGCCACCGGACCCAGTACAGCACTTCTCTACGTGGACCCCGTGCATCCTCATGAGCCCTCAAGTAA
CGTCCCTATGGCTGGCGGCTGACCTGACCTGACCGCAACTACCTGAACAGCGGCAACAGCCGGCACATCAGCATCGGCGGA
AGACTGTACCTGAGTCTGCTGCAACAGCATCTGGAATGCTACACCCCTGCCAGACCATAGCAGCCAGTTCGCGGTGAAGCT
GAAGATCGACTGGCCAAACCGGAAACAGCATCTTCAAGTACCGGGAAGATCCATCTGTACAGAGATCCACCCCAACAGAGCT
CATAGCGCGGCGAGCACCATCAGCCGGTGGGCAAGAACCTGAATAGCTGTCCGTGCCCGGATGTCATCAAGTGCAGCAGG
CCGGCAGAAACTTACCGGTGCCCTGCCAGCAGAAAGCAACAGCGAGATCATCTGTGCACCAACCCAGCCCTGCAGCAGTGAAC
CTGCAGTGGCCCTGAAACAAAGGCCCTTCTCATGCTGGACGGCATCTGAGCAAGTACTTGCACCTGATCTACGTGCACAAACCCG
TGTCAAGCCCTTCGAGAAGCCGCTGATGATCAGCATGGGCAATGAGAAGCTGTGGAAATCAAGGGCAACGACATCGACCCGAGG
CCGTGAAGGGGGAGGTGCTGAAAGTGGGCAACAAGTCTCGGAGAACATCCATCTGCACAGCGAGGCGGTGCTGTGCACCGTCC
AACGACCTGTGAAGCTGAACAGCGAGCTGAACATCGATGGAAGCAGGCCATCAGCAGCAGCGTGTGGGCAAGGATGTGTCAG
CCCGACAGAACTTACCGGCCCTGATCGCCGGGTGGTGTCTATCAGCACAGCCCTGCTGTGCTCCCTGGATTCTCTGTGGCTG
AAGAAGCGCAAGCAGATCAAGGATCTGGGCGAGCAACTGCTCAGATACGACGCCCGGGTGCACACCCCACTGGACAGACTGTT
GTCCCGCAGAAGCGTGTCCCACTACCGAGATGGTGTCCAACGAGAGCGTGGACTACCGGGCCACCTTCCCGGAGGACCACTTCC
CCAACCTCAGCCAGAAGCGGAGCTGCAGACAGGTGCAGTACCCCTGACCGACATGAGCCCACTCCTGACAGCGGCGACTCCGAC
ATCTCCAGCCCTGCTGAGAAATACCTGCACATCGACCTGAGCGCCCTGAACCCCGAGCTGGTGCAGGCGGTGCAGCAGCTGGT
AACGACGGCAAGAAATCCACTGGCCGCTGAAGTCCCTGAACCGGATCACCAGACATCGGCGAGGTGTCCAGTTTCTGACCGAGGGC
ATCATCATGAAGGACTTCAGCCACCCCAACGTGTGAGCCTGCTGGGCATCTGCTGAGAGCGAGGGCAGCCCTCTGGTGGTGTG
CCCTACATGAAGCAGCGCAGCTGCGGAACCTTATCCGGAACGAGACACAAACCCACCGTGAAGGATCTGATCGGATTCGGGTG
CAGTCCGCAAGGGATGAAGTATCTGGCCAGCAAGAAATCTGTCACCGCGACCTGGCCCGCAGAACTGCATGCTGGATGAGAAG
TTTACCGTGAAGTGGCCGACTTCCGGCTGGCCGGGATATGACGACAAAGAATATTACAGCTGCACAACAAGCAGCGCCCAAG
TGCCCGTGAAGTGGATGGCCCTGAAAGCCTGCAGACCCAGAAGTTCACCACAAAGTCCGAGCTGTGTCTTGGCGTGTGCTGT
GGGAGCTGATGACAGAGGCGCCCTCCCTACCCGACGTGAACACCTTCGACATACCCTGTACTCTGCAGGGCAGACGGCTG
CTGACGCCGAGTACTGCCCGACCCCTGTACGAAGTGTGCTGAAGTGTGGACCCCAAGGCGGAGATCGGGCCAGCTTCAG
CGAGTGTGTCGCCGATCAGCGCCATCTTCCACTTATCGGCGAGCAGTACTGTCAGTGCAGTGAACGCAACATACGTGAACGTGAAA
TCCGTGGCCCTACCTAGCTGTCTCAGCGAGGACAAAGCCGAGCAGAAAGTGGATACAGACCCGCGAGCTTCTGGGAGAC
AGCTAGCTTAGAAGATTCGAGAAATCGGACTGGCTTTCATTTCGACCCCATATGTGGAAGTCTGGGCGAGCGCATGCAGTACG
TCGATGTTGGTCCCGCGATGGCACCCCTGTGCTGTTCTGTCACGGTAACCCGACTCTCTCTACGTGTGGCGCAACATCATCCCGCA
TGTGACCGACCCATCGCTGCATTGCTCCAGACCTGATCGGTATGGGCAATCCGACAACCCAGACCTGGGTTATTTCTTCGACGAC
CACGTCCGCTTATGGATGCTTATCGAAGCCCTGGGTCTGGAAGAGTCTGCTGGTCAATCAGCAGTGGGCTCGCTCTGGGT
TTCCACTGGGCAAGCGCAATCCAGAGCGGTCAAAGGTATGCAATTTATGAGTTCATCCGCCCTATCCCGACCTGGGACGAATGGC
CAGAATTTGCCCGGAGACCTTCCAGGCCTTCCGACACCCGACGCTCGGCCGCAAGCTGATCATCGATCAGAACGTTTTTATCGAGG
GTACGCTGCCGATGGGTGCTGCCGCCGCTGACTGAAGTGCAGATGGACATACCGCGAGCCGTTCTGTAATCCTGTTGACCGCG
AGCCACTGTGGCGTTCCCAAAACGAGCTGCCAATCGCCGGTGAAGCAGCGAATCAGTCCGCGCTGGTGAAGAATACATGGACTGGC
TGCACAGTCCCTGTCCGAAGTGTGTTCTGGGGCACCCAGCGGTTCTGATCCACCGGCGAAGCCGCTCGCCCTGGCCAAA
AGCTCGCTAACTGAAGGCTGTGGACATCGGCCCGGCTGTAATCTGCTGCAAGAAGCAACCCGACCTGATCGGCGAGCGAGAT
CGCGCTGGCTGTCAGCGTGGAGATTCCGGC



ATGCCACCCTCCGGGACGGCCGGGGCAGCGCTCTGGCGTGTGGCTGGCGCTGCCCGGGAGTCGGGCTTGGAGGAAAAAGA
 AAGTTTGGCAAGGCACGAGTAACAAGCTCACGCAAGTTGGGCACCTTTGAAGATCATTTCACGCCTCCAGAGGATGTTCAATAACTGT
 GAGGTGGCTCTGGGAATTTGGAAATTACCTATGTGCAGAGGAATATGATCTTCTCTTAAAGACCATCCAGGAGGTGGCTGGTAT
 GTCCCTATTGGCCCTCAACACAGTGGAGCGAATTCCTTTGGAAAACCTGCAGATCATCAGAGGAATATGTACTACGAAAAATTCCTATGCC
 TTAGCAGTCTTATCTAACTATGATGCAATAAAAACCGGACTGAAGGAGTCCCATGAGAATTTACAGGAATCTGCATGGCCCGCTG
 CCGTTCCAGCAACACCTGCCCTGTGCACGTGGAGAGCATCCAGTGGCGGGACATAGTCAGCAGTGACTTTCAGCAACATGTGCA
 TGGACTTCCAGAACACCTGGGCAGCTGCCAAAAGTGTGATCCAAGCTGTCCAATGGGAGCTGCTGGGGTGCAGGAGAGGAGA
 GAACTGCCAGAACTGACCAAAATCATCTGTGCCAGCAGTGTCCGGGGCTGCCGTGGCAAGTCCCCAGTGACTGTGCCACAACAGT
 GTGCTGCAGGCTGCACAGGCCCCCGGGAGAGCGACTGCCTGGTCTGCCGCAAAATCCGAGACGAAGCCACGTGCAAGGACACCTGC
 CCCCCACTCATGCTCTACAACCCACACGTACAGATGGATGTGAACCCGAGGGGAAATACAGCTTTGGTGCCACCTGGCTGAAGA
 AGTGTCCCCGTAATTATGTGGTGACAGATCACGGCTCGTGGCTCCGAGCCTGTGGGGCCGACAGCTATGAGATGGAGGAAGACGGCG
 TCCGAAGTGTAAAGAGTGGCAAGGGCCCTTGGCCGAAAGTGTGAACGGAATAGGTATTGGTGAATTTAAAGACTCACTCTCCATAAT
 GCTACGAATATAAACACTTCAAAAACCTGCACCTCCATCAGTGGCGATCTCCACATCCTGCCGGTGGCAATTTAGGGGTGACTCTTAC
 ACATACTCTCTCTGGATCCACAGGAACCTGGATTTCTGAAAACCGTAAAGGAATCACAGGGTTTTTGTCTGATTGAGGCTTGGCCG
 AAAACAGGACGGACCTCCATGCCTTTGAGAACCTAGAAATCATACGGCGGAGGACCAAGCAACATGGTCACTGTTTCTCTGAGTCTG
 CAGCCTGAACATAACATCCTTGGGATTACGCTCCCTCAAGGAGATAAGTATGGAGATGTGATAATTCAGGAAACAAAAATTTGTGCTA
 TGCAATACATAAACTGAAAAAACTGTTGGGACCTCCGGTCAGAAAACCAAAATATAAGCAACAGAGGTGAAAAACAGCTGCAAGG
 CCACAGGCCAGGTCTGCCATGCTTGTGCTCCCGAGGGCTGTGGGGCCCGGAGCCAGGGACTGCCTCTTGGCCGAATGTG
 AGCCGAGGCAGGGAATGCTGGACAAAGTGAACCTTCTGGAGGTGAGCCAAAGGAGTTTGGGAGAACTCTGAGTGCATACAGTGC
 CACCCAGAGTGCCTGCCTCAGGCCATGAACATCACCTGCACAGGACGGGGACAGACAACCTGATCCAGTGTGCCACTACATTTGAGC
 GCCCCACTGCGTCAAGACCTGCCCGGCAGGAGTCAAGGAGAAAACAAACCCCTGGTCTGGAAGTACCGAGACCGCCGGCCATGTG
 TGCCACTGTGCCATCCAACTGCACCTAGCGATGCACTGGGCCAGGTCTTGAAGGCTGTCCAGCAATGGGCTAAGATCCCGTCCCA
 TCGCACTGGGATGGTGGGGCCCTCTCTGTGCTGGTGGTGGCCCTGGGATCGGCCTCTCATGCCAAGCGCCACATCGTT
 CGGAAGCGCAGCTGGGAGGCTGTGCAGGAGAGGGAGCTTGTGGAGCTCTTACACCCAGTGAGAGCTCCCAACCAAGCTCT
 CTTGAGGATCTTGAAGAAAATGAATTCAAAAGATCAAAGTGTGGCTCCGGTGCCTTCCGACCGGTGATAAGGACTCTGGATC
 CCAGAAAGTGGAGAAAGTTAAATTCCTGCTGCTATCAAGGAATTAAGAGAAAGCAACATCTCCGAAAGCCAAAGGAATCTCGATGA
 AGCTTACGTGATGGCCAGCGTGGACAACCCACAGTGTGCCGCTGTGGGATCTGCTCCTCCTCCAGTGCAGCTCATCACGCA
 GCTCATGCCCTTCGGCTGCCTCTGGACTATGTCGGGAACACAAAGACAATATTGGCTCCAGTACCTGCTCAACTGGTGTGTGAG
 ATCCGAAAGGGCATGAAGTACTTGGAGGACCGTCCGCTTGGTGCACCGCAGCTTGGCAGCCAGGAACCTACTGGTAAAAACACCGCAG
 CATGTCAAGATCACAGATTTTGGGCTGGCCAACTGCTGGTGGCGGAAGAGAAAGAAATACCATGCAGAAGGAGGCAAGTCCATCA
 AGTGGATGGCATTGGAATCAATTTTACACAGAATCTATCCACAGAGTGTGATGCTGGAGCTACGGGGTGGCCGTTTGGGAGTGTGAT
 ACCTTTGGATCCAAGCCATATGACGGAATCCCTGCCAGCGAGATCTCCTCCATCTGGAGAAAGGAGAACCGCTCCCTCAGCCACCCA
 TATGTACCATCGATGCTACATGATCATGGTCAAGTGTGGATGATAGACGAGATAGTCGCCCAAAAGTTCCGTGAGTTGATCATGAAT
 CTCAAAATGGCCCGAGACCCCGCCTACCTTGTCAATTCAGGGGGATGAAAGATGCAATTTGCCAAGTCTCACAGACTCCAACCTC
 TACCGTGCCTGATGGATGAAAGACATGGACGACGCTGGTGGATGCCGACGAGTACCTCATCCACAGCAGGGCTTCTTACGACG
 CCCTCCAGTCAAGGACTCCCTCCTGAGCTCTGAGTGAACAGCAACATTCACCGTGGCTTGCATTGATAGAAATGGGCTGC
 AAAGTGTCCATCAAGGAAGACAGCTTCTGAGCGATACAGCTCAGACCCACAGGCGCCTTACTGAGGACAGCATAGACGACAC
 CTTCTCCAGTGCCTGAATACATAAACAGTCCCTTCCAAAAGGCCCGCTGGCTGTGTCAGAACTCTGTCTATCACAATCAGCCTC
 TGAACCCCGCCCGCAGCAGACCCACACTACCAGGACCCACAGCACTGCAGTGGCAACCCCGAGTATCTCAACACTGTCCAGC
 CCACCTGTGTCAACAGCACATTCGACAGCCCTGCCACTGGGGCCAGAAAGGCAGCCACCAATTAGCCTGGACAACCTGACTACC
 AGCAGGACTTCTTCCAAAGGAAGCCAAAGCAATGGATCTTTAAGGCTCCACAGCTGAAATGCAGAAATCAAGGGTCCGGCC
 ACAAAGCAGTGAATTTATGGAGCAATTTGGACAAGACTGCGAAATGAAGCCACCCCTGGATAGCCCTTGGGCAAGCTGGAA
 CTGTCTGGTGGCAAGGGCTGCACCTATCATCTTCTGGGCAAGGAACATCTCCGCGGAGCCGCTGGAGTGGCTGCCCA
 GCCGCGTGTGGGGGACAGGCCACTGATGCAGGCCACCGCTGGCTCAACGCTTACTTCAACAGCTGAGGCCATCGAGGA
 GTTCCCTGTGCCAGCCCTGCACACCCAGTGTCCAGCAGGAGAGCTTACCAGCCAGGTGCTGTGAAACTGCTGAAAGTGGTGA
 GTTCGGAGAGGTATCAGCTACAGCCACCTGGCCGCTGGCCGCAATCCCGCCACCGCCGCTGAAACCCGCTGAGC
 GGAAATCCCGTGCCTTCTGATCCCTGCCACCGGGTGGTGCAGGGCGACTGGAGCTGGGGGGCTCAGGCGGGCTCGCCG
 TGAAGAGTGGCTGTGGCCACGAGGGCCAGACTGGGCAAGCTGGGCTGGTCAATTGCACCATCACCATACCAC

Abbreviations

Akt	Protein kinase B
BG	Benzylguanine
c-Met	Hepatocyte growth factor receptor
CP	Benzylchloropyrimidine
DhaA	Haloalkane dehalogenase
DMF	Dimethylformamid
Dipea	N,N-diisopropylethylamine
DNA	Deoxyribonucleic acid
DOL	Degree of labeling
eGFP	Enhanced green fluorescent protein
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
FRET	Förster resonance energy transfer
gSEP	GFP-enhanced staining efficiency probe
hAGT	O6-alkylguanine-DNA alkyltransferase
HGF	Hepatocyte growth factor
HMM	Hidden Markov model
IFN	Interferon
IFNAR1	Interferon- α receptor 1
IFNAR2	Interferon- α receptor 2
LynG	Lyn kinase membrane anchor GFP fusion
MAPK	Mitogen activated protein kinase
NSCLC	Non-small cell lung cancer
PCR	Polymerase chain reaction
PFA	Paraformaldehyde
PI3K	3-phosphoinositide-dependent protein kinase
PIP3	phosphatidylinositol-3,4,5-trisphosphate
PTB	Phosphotyrosine binding
ROI	Region of interest
SH2	Src homology 2
SiR	Siliconrhodamine
smALEX	Single-molecule alternating laser excitation
TIRFM	Total internal reflection fluorescence microscopy
TKI	Tyrosine kinase inhibitor
TMR	Tetramethylrhodamine

Bibliography

- [1] Tsien, R. Y. *Annual Review of Biochemistry* **1998**, *67*, 509–544.
- [2] Shimomura, O.; Johnson, F. H.; Saiga, Y. *Journal of Cellular and Comparative Physiology* **1962**, *59*, 223–239.
- [3] Morise, H.; Shimomura, O.; Johnson, F. H.; Winant, J. *Biochemistry* **1974**, *13*, 2656–2662.
- [4] Giepmans, B. N. G. *Science* **2006**, *312*, 217–224.
- [5] Cranfill, P. J.; Sell, B. R.; Baird, M. A.; Allen, J. R.; Lavagnino, Z.; de Gruiter, H. M.; Kremers, G.-J.; Davidson, M. W.; Ustione, A.; Piston, D. W. *Nature Methods* **2016**, .
- [6] Resch-Genger, U.; Grabolle, M.; Cavaliere-Jaricot, S.; Nitschke, R.; Nann, T. *Nature Methods* **2008**, *5*, 763–775.
- [7] Medintz, I. L.; Uyeda, H. T.; Goldman, E. R.; Mattoussi, H. *Nature Materials* **2005**, *4*, 435–446.
- [8] Kim, C. H.; Axup, J. Y.; Schultz, P. G. *Current Opinion in Chemical Biology* **2013**, *17*, 412–419.
- [9] Crivat, G.; Taraska, J. W. *Trends in Biotechnology* **2012**, *30*, 8–16.
- [10] Lakowicz, J. R. *Principles of fluorescence spectroscopy*; Springer: New York, 3rd ed ed.; 2006.
- [11] Roy, R.; Hohng, S.; Ha, T. *Nature Methods* **2008**, *5*, 507–516.
- [12] Nobel Media AB 2014, “The Nobel Prize in Chemistry 2014”, http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2014/, Accessed: 2017-06-12.
- [13] Huisken, J. *Science* **2004**, *305*, 1007–1009.
- [14] Balzarotti, F.; Eilers, Y.; Gwosch, K. C.; GynnÄŸ, A. H.; Westphal, V.; Stefani, F. D.; Elf, J.; Hell, S. W. *Science* **2017**, *355*, 606–612.
- [15] Presman, D. M.; Ball, D. A.; Paakinaho, V.; Grimm, J. B.; Lavis, L. D.; Karpova, T. S.; Hager, G. L. *Methods* **2017**, .
- [16] ATTO-TEC GmbH, “Properties of ATTO-Dyes”, http://www.atto-tec.com/fileadmin/user_upload/Katalog_Flyer_Support/Dye_Properties_01.pdf, Accessed: 2017-06-12.

- [17] ThermoFisher Scientific, “Alexa Fluor Dyes-Across the Spectrum”, <https://www.thermofisher.com/de/de/home/brands/molecular-probes/key-molecular-probes-products/alexa-fluor/alexa-fluor-dyes-across-the-spectrum.html>, Accessed: 2017-06-12.
- [18] Großmayer, K. S.; Kurz, A.; Herten, D.-P. *ChemPhysChem* **2014**, *15*, 734–742.
- [19] Mitronova, G. Y.; Polyakova, S.; Wurm, C. A.; Kolmakov, K.; Wolfram, T.; Meineke, D. N. H.; Belov, V. N.; John, M.; Hell, S. W. *European Journal of Organic Chemistry* **2015**, *2015*, 337–349.
- [20] Juillerat, A.; Gronemeyer, T.; Keppler, A.; Gendreizig, S.; Pick, H.; Vogel, H.; Johnsson, K. *Chemistry & biology* **2003**, *10*, 313–317.
- [21] Moschel, R. C.; McDougall, M. G.; Dolan, M. E.; Stine, L.; Pegg, A. E. *Journal of medicinal chemistry* **1992**, *35*, 4486–4491.
- [22] Gronemeyer, T.; Chidley, C.; Juillerat, A.; Heinis, C.; Johnsson, K. *Protein Engineering Design and Selection* **2006**, *19*, 309–316.
- [23] Sun, X.; Zhang, A.; Baker, B.; Sun, L.; Howard, A.; Buswell, J.; Maurel, D.; Masharina, A.; Johnsson, K.; Noren, C. J.; Xu, M.-Q.; Corrêa, I. R. *ChemBioChem* **2011**, *12*, 2217–2226.
- [24] Bosch, P.; Corrêa, I.; Sonntag, M.; Ibach, J.; Brunsveld, L.; Kanger, J.; Subramaniam, V. *Biophysical Journal* **2014**, *107*, 803–814.
- [25] Los, G. V. *et al.* *ACS Chemical Biology* **2008**, *3*, 373–382.
- [26] Urh, M. *Current Chemical Genomics* **2013**, *6*, 72–78.
- [27] Stöhr, K.; Sieberg, D.; Ehrhard, T.; Lymperopoulos, K.; Öz, S.; Schulmeister, S.; Pfeifer, A. C.; Bachmann, J.; Klingmüller, U.; Sourjik, V.; Herten, D.-P. *Analytical Chemistry* **2010**, *82*, 8186–8193.
- [28] Liu, Y.-C. C.; Rieben, N.; Iversen, L.; Sørensen, B. S.; Park, J.; Nygård, J.; Martinez, K. L. *Nanotechnology* **2010**, *21*, 245105.
- [29] Roder, F.; Wilmes, S.; Richter, C. P.; Piehler, J. *ACS Chemical Biology* **2014**, 2479–2484.
- [30] Wilmes, S.; Beutel, O.; Li, Z.; Francois-Newton, V.; Richter, C. P.; Janning, D.; Kroll, C.; Hanhart, P.; Hotte, K.; You, C.; Uze, G.; Pellegrini, S.; Piehler, J. *The Journal of Cell Biology* **2015**, *209*, 579–593.
- [31] Latty, S.; Felce, J.; Weimann, L.; Lee, S.; Davis, S.; Klenerman, D. *Biophysical Journal* **2015**, *109*, 1798–1806.
- [32] Cognet, L.; Leduc, C.; Lounis, B. *Current Opinion in Chemical Biology* **2014**, *20*, 78–85.
- [33] Brandenburg, B.; Zhuang, X. *Nature Reviews Microbiology* **2007**, *5*, 197–208.

- [34] Fujiwara, T.; Ritchie, K.; Murakoshi, H.; Jacobson, K.; Kusumi, A. *The Journal of Cell Biology* **2002**, *157*, 1071–1082.
- [35] Low-Nam, S. T.; Lidke, K. A.; Cutler, P. J.; Roovers, R. C.; van Bergen en Henegouwen, P. M. P.; Wilson, B. S.; Lidke, D. S. *Nature Structural & Molecular Biology* **2011**, *18*, 1244–1249.
- [36] Jaqaman, K.; Loerke, D.; Mettlen, M.; Kuwata, H.; Grinstein, S.; Schmid, S. L.; Danuser, G. *Nature Methods* **2008**, *5*, 695–702.
- [37] Tinevez, J.-Y.; Perry, N.; Schindelin, J.; Hoopes, G. M.; Reynolds, G. D.; Laplantine, E.; Bednarek, S. Y.; Shorte, S. L.; Eliceiri, K. W. *Methods (San Diego, Calif.)* **2017**, *115*, 80–90.
- [38] Chenouard, N. *et al. Nature Methods* **2014**, *11*, 281–289.
- [39] Hern, J. A.; Baig, A. H.; Mashanov, G. I.; Birdsall, B.; Corrie, J. E. T.; Lazareno, S.; Molloy, J. E.; Birdsall, N. J. M. *Proceedings of the National Academy of Sciences* **2010**, *107*, 2693–2698.
- [40] Valley, C. C.; Arndt-Jovin, D. J.; Karedla, N.; Steinkamp, M. P.; Chizhik, A. I.; Hlavacek, W. S.; Wilson, B. S.; Lidke, K. A.; Lidke, D. S. *Molecular biology of the cell* **2015**, *26*, 4087–4099.
- [41] Borden, E. C.; Sen, G. C.; Uze, G.; Silverman, R. H.; Ransohoff, R. M.; Foster, G. R.; Stark, G. R. *Nature Reviews Drug Discovery* **2007**, *6*, 975–990.
- [42] Piehler, J.; Thomas, C.; Garcia, K. C.; Schreiber, G. *Immunological reviews* **2012**, *250*, 317–334.
- [43] Thomas, C.; Moraga, I.; Levin, D.; Krutzik, P.; Podoplelova, Y.; Trejo, A.; Lee, C.; Yarden, G.; Vleck, S.; Glenn, J.; Nolan, G.; Piehler, J.; Schreiber, G.; Garcia, K. *Cell* **2011**, *146*, 621–632.
- [44] Stark, G. R.; Kerr, I. M.; Williams, B. R. G.; Silverman, R. H.; Schreiber, R. D. *Annual Review of Biochemistry* **1998**, *67*, 227–264.
- [45] Brierley, M. M.; Fish, E. N. *Journal of interferon & cytokine research* **2002**, *22*, 835–845.
- [46] Li, X.; Leung, S.; Kerr, I. M.; Stark, G. R. *Molecular and cellular biology* **1997**, *17*, 2048–2056.
- [47] Löchte, S.; Waichman, S.; Beutel, O.; You, C.; Piehler, J. *The Journal of Cell Biology* **2014**, *207*, 407–418.
- [48] Jaks, E.; Gavutis, M.; Uzé, G.; Martal, J.; Piehler, J. *Journal of Molecular Biology* **2007**, *366*, 525–539.
- [49] Levin, D.; Harari, D.; Schreiber, G. *Molecular and Cellular Biology* **2011**, *31*, 3252–3266.
- [50] Torre, L. A.; Bray, F.; Siegel, R. L.; Ferlay, J.; Lortet-Tieulent, J.; Jemal, A. *CA: A Cancer Journal for Clinicians* **2015**, *65*, 87–108.

- [51] Lawrence, M. S. *et al. Nature* **2013**, *499*, 214–218.
- [52] Cui, J.; Hu, Y.-F.; Feng, X.-M.; Tian, T.; Guo, Y.-H.; Ma, J.-W.; Nan, K.-J.; Zhang, H.-Y. *Tumor Biology* **2014**, *35*, 11701–11709.
- [53] Park, S.; Langley, E.; Sun, J.-M.; Lockton, S.; Ahn, J. S.; Jain, A.; Park, K.; Singh, S.; Kim, P.; Ahn, M.-J. *Oncotarget* **2015**, *6*, 30929–30938.
- [54] Olayioye, M. A. *The EMBO Journal* **2000**, *19*, 3159–3167.
- [55] Macdonald-Obermann, J. L.; Pike, L. J. *Journal of Biological Chemistry* **2014**, *289*, 26178–26188.
- [56] Birchmeier, C.; Birchmeier, W.; Gherardi, E.; Vande Woude, G. F. *Nature Reviews Molecular Cell Biology* **2003**, *4*, 915–925.
- [57] Tao, R.-H.; Maruyama, I. N. *Journal of Cell Science* **2008**, *121*, 3207–3217.
- [58] Alberts, B., Ed.; *Molecular biology of the cell*; Garland Science: New York, 5th ed.; 2008.
- [59] Zhang, W.; Liu, H. T. *Cell Research* **2002**, *12*, 9–18.
- [60] Stewart, E. L.; Tan, S. Z.; Liu, G.; Tsao, M.-S. *Translational lung cancer research* **2014**, *4*, 67–81.
- [61] Blakely, C. M.; Bivona, T. G. *Cancer Discovery* **2012**, *2*, 872–875.
- [62] Fujino, S.; Enokibori, T.; Tezuka, N.; Asada, Y.; Inoue, S.; Kato, H.; Mori, A. *European Journal of Cancer* **1996**, *32*, 2070–2074.
- [63] Reck, M.; Popat, S.; Reinmuth, N.; De Ruyscher, D.; Kerr, K. M.; Peters, S.; on behalf of the ESMO Guidelines Working Group, *Annals of Oncology* **2014**, *25*, iii27–iii39.
- [64] Solca, F.; Dahl, G.; Zoephel, A.; Bader, G.; Sanderson, M.; Klein, C.; Kraemer, O.; Himmelsbach, F.; Haaksma, E.; Adolf, G. R. *Journal of Pharmacology and Experimental Therapeutics* **2012**, *343*, 342–350.
- [65] Cross, D. A. E. *et al. Cancer Discovery* **2014**, *4*, 1046–1061.
- [66] Yu, H. A.; Arcila, M. E.; Rekhman, N.; Sima, C. S.; Zakowski, M. F.; Pao, W.; Kris, M. G.; Miller, V. A.; Ladanyi, M.; Riely, G. J. *Clinical Cancer Research* **2013**, *19*, 2240–2247.
- [67] Yun, C.-H.; Mengwasser, K. E.; Toms, A. V.; Woo, M. S.; Greulich, H.; Wong, K.-K.; Meyerson, M.; Eck, M. J. *Proceedings of the National Academy of Sciences* **2008**, *105*, 2070–2075.
- [68] Thress, K. S. *et al. Nature Medicine* **2015**, *21*, 560–562.
- [69] Chen, H.; Cheng, X. *OncoTargets and Therapy* **2014**, 1689–1704.
- [70] Tanizaki, J.; Okamoto, I.; Sakai, K.; Nakagawa, K. *British Journal of Cancer* **2011**, *105*, 807–813.

- [71] Fichter, C. D.; Gudernatsch, V.; Przypadlo, C. M.; Follo, M.; Schmidt, G.; Werner, M.; Lassmann, S. *Journal of Molecular Medicine* **2014**, *92*, 1209–1223.
- [72] Moores, S. L. *et al. Cancer Research* **2016**, *76*, 3942–3953.
- [73] Arai, R.; Ueda, H.; Kitayama, A.; Kamiya, N.; Nagamune, T. *Protein engineering* **2001**, *14*, 529–532.
- [74] Resh, M. D. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* **1999**, *1451*, 1–16.
- [75] Gomez, M.; Jansen, L. E. “based pulse labeling for analysis of protein turnover in living cells”, <http://www.neb.com/products/~media/Catalog/All-Products/34A20E471BoC4C5D85369BAoA1212E53/Application%20Notes/appNoteS9106.pdf>, urldate: 2015-07-31.
- [76] Ketteler, R.; Glaser, S.; Sandra, O.; Martens, U. M.; Klingmüller, U. *Gene Therapy* **2002**, *9*, 477–487.
- [77] Morgenstern, J. P.; Land, H. *Nucleic Acids Research* **1990**, *18*, 3587–3596.
- [78] Schindelin, J. *et al. Nature Methods* **2012**, *9*, 676–682.
- [79] de Chaumont, F.; Dallongeville, S.; Chenouard, N.; Hervé, N.; Pop, S.; Provoost, T.; Meas-Yedid, V.; Pankajakshan, P.; Lecomte, T.; Le Montagner, Y.; Lagache, T.; Dufour, A.; Olivo-Marin, J.-C. *Nature Methods* **2012**, *9*, 690–696.
- [80] You, C.; Wilmes, S.; Richter, C. P.; Beutel, O.; Liße, D.; Piehler, J. *ACS Chemical Biology* **2013**, *8*, 320–326.
- [81] Kiel, A.; Kovacs, J.; Mokhir, A.; Krämer, R.; Herten, D.-P. *Angewandte Chemie International Edition* **2007**, *46*, 3363–3366.
- [82] Hodak, J. H.; Downey, C. D.; Fiore, J. L.; Pardi, A.; Nesbitt, D. J. *Proceedings of the National Academy of Sciences* **2005**, *102*, 10505–10510.
- [83] Kuno, M.; Fromm, D. P.; Hamann, H. F.; Gallagher, A.; Nesbitt, D. J. *The Journal of Chemical Physics* **2001**, *115*, 1028–1040.
- [84] Maurel, D.; Comps-Agrar, L.; Brock, C.; Rives, M.-L.; Bourrier, E.; Ayoub, M. A.; Bazin, H.; Tinel, N.; Durroux, T.; Prézeau, L.; Trinquet, E.; Pin, J.-P. *Nature Methods* **2008**, *5*, 561–567.
- [85] Lukinavičius, G. *et al. Nature Chemistry* **2013**, *5*, 132–139.
- [86] Klein, T.; Löscherger, A.; Proppert, S.; Wolter, S.; van de Linde, S.; Sauer, M. *Nature Methods* **2011**, *8*, 7–9.
- [87] Calebiro, D.; Rieken, F.; Wagner, J.; Sungkaworn, T.; Zabel, U.; Borzi, A.; Cocucci, E.; Zurn, A.; Lohse, M. J. *Proceedings of the National Academy of Sciences* **2013**, *110*, 743–748.
- [88] Ulbrich, M. H.; Isacoff, E. Y. *Nature Methods* **2007**, 319–321.

- [89] Foo, Y.; Naredi-Rainer, N.; Lamb, D.; Ahmed, S.; Wohland, T. *Biophysical Journal* **2012**, *102*, 1174–1183.
- [90] Finan, K.; Raulf, A.; Heilemann, M. *Angewandte Chemie International Edition* **2015**, *54*, 12049–12052.
- [91] You, C.; Marquez-Lago, T. T.; Richter, C. P.; Wilmes, S.; Moraga, I.; Garcia, K. C.; Leier, A.; Piehler, J. *Science Advances* **2016**, *2*, e1600452–e1600452.
- [92] Bag, N.; Yap, D. H. X.; Wohland, T. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2014**, *1838*, 802–813.
- [93] Saha, S.; Lee, I.-H.; Polley, A.; Groves, J. T.; Rao, M.; Mayor, S. *Molecular biology of the cell* **2015**, *26*, 4033–4045.
- [94] ThermoFisher Scientific, “Phusion Hot Start II High-Fidelity DNA Polymerase, Product Information”, https://tools.thermofisher.com/content/sfs/manuals/MAN0012401_Phusion_HotStartII_DNAPolymerase_500U_UG.pdf, Accessed: 2016-11-22.
- [95] New England Biolabs, “NEBuffer Activity/Performance Chart with Restriction Enzymes”, <https://www.neb.com/tools-and-resources/usage-guidelines/nebuffer-performance-chart-with-restriction-enzymes?device=pdf>, Accessed: 2016-11-22.
- [96] Qiagen, “QIAquick PCR Purification Kit, Quick-Start Protocol”, <https://www.qiagen.com/us/resources/download.aspx?id=390a728a-e6fc-43f7-bf59-b12091cc4380&lang=en>, Accessed: 2016-11-22.
- [97] Qiagen, “QIAEX II Gel Extraction Kit , Quick-Start Protocol”, <https://www.qiagen.com/us/resources/download.aspx?id=8a1e1b5f-70be-480b-bc9d-8eb444ff7b63&lang=en>, Accessed: 2016-11-22.
- [98] New England Biolabs, “Quick Ligation Protocol (M2200)”, <https://www.neb.com/protocols/1/01/01/quick-ligation-protocol>, Accessed: 2016-11-22.
- [99] Qiagen, “QIAprep Spin Miniprep Kit, Quick-Start Protocol”, <https://www.qiagen.com/gb/resources/download.aspx?id=331740ca-077f-4ddd-9e5a-2083f98eebd5&lang=en>, Accessed: 2016-11-22.
- [100] Genomed, “Jetstar 2.0 Mini, Midi, Maxi Protocol”, <http://www.genomed-dna.com/pdf/Star-PDFs/JETSTAR-Protocol.PDF>, Accessed: 2016-11-22.
- [101] Invitrogen, “PureLink HiPure Plasmid DNA Purification Kits, Quick Reference”, https://tools.thermofisher.com/content/sfs/manuals/purelink_hipure_plasmid_qrc.pdf, Accessed: 2016-11-22.
- [102] Nolan, G. L. “Phoenix helper-free retrovirus producer lines”, https://web.stanford.edu/group/nolan/_OldWebsite/retroviral_systems/phx.html, Accessed: 2016-11-22.
- [103] Graz, T. U. “Characteristics of 5-TAMRA at pH 7”, <http://www.fluorophores.tugraz.at/substance/171>, Accessed: 2017-01-13.

-
- [104] Spirochrome, "Ex/Em spectra of SiR-NHS ester", http://spirochrome.com/documents/SiR_abs_em.xls, Accessed: 2017-01-13.

Publications

Journal Articles

1. S. Hänselmann, D.-P. Hertel "Two-Color Single-Molecule Tracking in Live Cells" (2017). In: H. Erfle (Ed.): *Super-Resolution Microscopy: Methods and Protocols*, Humana Press, DOI: 10.1007/978-1-4939-7265-4.
2. S. Hänselmann, D.-P. Hertel "Single-Molecule Spectroscopy" (2017). In: J. C. Lindon, G. E. Tranter, D. W. Koppenaal (Eds.): *Encyclopedia of Spectroscopy and Spectrometry* 4, Elsevier, 84-88. DOI: 10.1016/B978-0-12-409547-2.12093-1.

Conference Poster Contributions

1. *PicoQuant 20th International Workshop on "Single Molecule Spectroscopy and Ultra-sensitive Analysis in the Life Sciences"*, Berlin 2014, Poster: "Type I interferon signaling: Towards kinetic and stoichiometric parameters of the signaling complex."
2. *Labeling & Nanoscopy*, Heidelberg 2015, Poster: "Type I interferon signaling: Towards kinetic and stoichiometric parameters of the signaling complex."
3. *114th General Assembly of the German Bunsen Society for Physical Chemistry*, Bochum 2015, Poster: "Measuring the kinetics and stoichiometry of the Type I interferon signaling complex."
4. *PicoQuant 22th International Workshop on "Single Molecule Spectroscopy and Ultra-sensitive Analysis in the Life Sciences"*, Berlin 2016, Poster: "Measuring the kinetics and stoichiometry of protein complex formation on a single-molecule level."

Acknowledgments

Ich möchte mich bei allen bedanken, die diese Doktorarbeit ermöglicht und zu ihrem Gelingen beigetragen haben.

Bei meinem Betreuer Dirk-Peter Herten für die großartige fachliche, persönliche und administrative Unterstützung und für die Freiheit, die ich bei der Bearbeitung des Themas genießen durfte. Bei Ursula Kummer für die Übernahme der Erstbetreuerschaft und dafür, dass sie immer ein offenes Ohr für mich hatte. Bei Ursula Klingmüller für die uneingeschränkte Nutzung ihres Labors. Ohne die Unterstützung ihrer Gruppe wäre diese Arbeit undenkbar gewesen. Allen Dreien gilt ein besonderer Dank dafür, dass sie mir als TAC Member dabei geholfen haben, die Richtung nicht zu verlieren.

Ganz sicher wäre diese Arbeit nicht ohne die Unterstützung von Florian Salopiata entstanden. Vielen Dank, dass ich dein Molekularbiologie-Padawan sein durfte und du mir etliche Western Blots entwickelt und stabile Zelllinien gebaut hast. Die Zusammenarbeit hat mir viel Spaß gemacht! Bei Frederique Kok möchte ich mich dafür bedanken, dass sie auf dem steinigen Weg der CRISP/Cas9 Knock-Ins immer wieder die Initiative ergriffen hat. Auch wenn das Projekt schließlich keinen Erfolg hatte, habe ich viel über genome editing gelernt und meine Frustrationstoleranz erheblich ausgebaut.

Bei Nikolas Schnellbacher und Sebastian Malkusch für die stressfreie Zusammenarbeit und bei Yu Qiao für seine Ausdauer.

Allen Mitgliedern des "Immunoquant" Konsortiums, die mir geholfen haben, meine Arbeit in einem breiteren Kontext zu sehen. Dem Bundesministerium für Bildung und Forschung sowie der FAZIT Stiftung für die Finanzierung dieser Doktorarbeit. Und natürlich Angelika Neuner, dafür dass sie, in einer unvergleichlich freundlichen Art, immer den Überblick im Bürokratie- und Finanzjungle behalten hat.

Der ganzen Arbeitsgruppe für die angenehme Atmosphäre während und neben der Arbeit! Anne Unger besonders für die Einführung in Zellkultur und Proteinlabeling, Daniel Barzan für Erdbeermargaritas, Arina Rybina für russische Süßigkeiten, Kristin Grußmayer für stetige konstruktive Kritik, Dominik Brox dafür, dass ich ihn immer alles fragen konnte, Andreas Haderspeck für Gespräche über Kaffee, Urlaub und Essen, Ilse Valtierra für die death-by-chocolate brownies, Sebastian Bierbaum für den guten Kick, Klaus Yserentant für die Lösung aller Netzwerkprobleme, Christopher Schmid für guten Kuchen, Martin Seefeld für seinen unabsichtlichen Humor, Wioleta Chmielewicz für polnischen Wodka und Felix Braun für das Aufhübschen meiner rudimentären Matlab-Codes.

Ein sehr großer Dank geht natürlich an meine Eltern, die mir diesen Weg ermöglicht haben und auf deren Unterstützung ich immer zählen kann. Zuletzt möchte ich meiner Freundin Sarah danken, ohne die ich nicht wäre, wer ich bin. Danke, dass du mich schon so lange aushältst und mein Leben wundervoll machst. Du bist ein Teil meiner Geschichte!

Eidesstattliche Versicherung

gemäß § 8 der Promotionsordnung der
Naturwissenschaftlich- Mathematischen Gesamtfakultät der
Universität Heidelberg

1. Bei der eingereichten Dissertation zu dem Thema “**Quantitative Microscopy: Measuring membrane receptor interactions in live cells**” handelt es sich um meine eigenständig erbrachte Leistung.
2. Ich habe nur die angegebenen Quellen und Hilfsmittel benutzt und mich keiner unzulässigen Hilfe Dritter bedient. Insbesondere habe ich wörtlich oder sinngemäß aus anderen Werken übernommene Inhalte als solche kenntlich gemacht.
3. Die Arbeit oder Teile davon habe ich bislang nicht an einer Hochschule des In- oder Auslands als Bestandteil einer Prüfungs- oder Qualifikationsleistung vorgelegt.
4. Die Richtigkeit der vorstehenden Erklärungen bestätige ich.
5. Die Bedeutung der eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Versicherung sind mir bekannt.

Ich versichere an Eides statt, dass ich nach bestem Wissen die reine Wahrheit erkläre und nichts verschwiegen habe.

Heidelberg, den 18. Juli 2017

Siegfried Hänselmann