

# Optical imaging methods for the study of disease models from the nano to the mesoscale



**Pedro Pablo Vallejo Ramirez**

Department of Chemical Engineering and Biotechnology  
University of Cambridge

This dissertation is submitted for the degree of  
*Doctor of Philosophy*



*To my parents, my sister, and my brother,*

*“En algún anaquel de algún hexágono....”*

*(“On some shelf of some hexagon....”)*

- Jorge Luis Borges, The Library of Babel





## **Declaration**

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Acknowledgements and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. This thesis contains fewer than 65,000 words including appendices, bibliography, footnotes, tables and equations and has fewer than 150 figures as specified by the Engineering Degree Committee.

Pedro Pablo Vallejo Ramirez  
July 2020



# **Optical imaging methods for the study of disease models from the nano to the mesoscale**

**Pedro Pablo Vallejo Ramirez**

## **Abstract**

The visualisation of disease phenotypes allows scientists to study fundamental mechanisms of disease. Optical imaging methods are useful not only to observe anatomical features of biological samples, but also to infer interactions between molecular species using fluorescence labelling. This thesis presents the development of imaging and analysis tools to study biological questions in three models of disease, with samples ranging from the sub-cellular to the organ scale.

First, the role of the alpha-synuclein (a-syn) protein, whose dysfunction is a hallmark of Parkinson's Disease, was studied with respect to vesicle trafficking at the synapse. Synaptic vesicles are ~40 nm in diameter; imaging vesicles therefore requires methods with resolution below the diffraction limit. Single-molecule localisation microscopy (SMLM), which circumvents the diffraction limit by separating fluorophore emission in time to localise individual molecules in space with ~20 nm precision, was thus implemented to study a-syn in purified synaptic boutons. A software package was developed to analyse the colocalisation of a-syn with internalised vesicles, and the clustering of a-syn under differing synaptic calcium levels. The colocalisation of a-syn and internalised vesicles was found to be temperature independent, suggesting that a-syn is involved in non-canonical trafficking mechanisms. Ground truth simulations from a synaptosome model were used to benchmark two cluster analysis methods. Both methods applied on the experimental data showed that a-syn becomes less clustered at low synaptic calcium levels.

Second, the spatiotemporal association of ESCRT-II, a protein complex whose role in the budding of the human immunodeficiency virus (HIV) was previously considered dispensable, and the HIV polyprotein Gag was studied during viral egress using novel image analysis tools. A nearest-neighbour analysis showed the ESCRT-II protein EAP45 colocalises with Gag similarly to ALIX, a protein well known to be involved in HIV budding. However, upon deletion of EAP45's N-terminus, its colocalisation with Gag was significantly impaired, highlighting the importance of this EAP45 domain in linking to Gag. Single particle tracking was used to trace the trajectories of EAP45 and Gag in live cells, and an algorithm was developed to visualise the simultaneous motion of two particles; these analyses revealed three types of potential dynamic interaction between EAP45 and Gag.

Finally, an open-source instrument to visualise phenotypes from large organs in 3D was developed for the study of chronic obstructive pulmonary disease (COPD) models. The instrument implements Optical Projection Tomography, a technique which can reconstruct cross-sectional slices of a transparent object from its orthographic projections, using off-the-shelf components and novel ImageJ plugins for artefact correction and volume reconstructions. Excised and cleared mouse lungs were imaged in which high order airways can be discerned with 50  $\mu\text{m}$  resolution. The raw lung data, instructions for building the instrument, the free ImageJ plugins, and a detailed software manual are available in an online repository to encourage the widespread use of OPT for imaging large samples.

## Acknowledgements

My deepest gratitude goes out to:

*Prof. Clemens Kaminski*, for welcoming me into his lab, and teaching me how to craft stories and see the big picture in scientific projects,

*Dr. Eric J Rees*, for his kind day-to-day mentorship, and for showing me how to deliver an engaging scientific talk,

*Dr. Gabi Kaminski*, for supporting my curiosity to study neurodegeneration,

*Dr. Steven Lee*, for the cascades of ideas and sage advice during our meetings,

The *Gates Cambridge Trust*, for providing a welcoming community of extraordinary scholars,

*Prof. Wayne Knox* and *Prof. Julie Bentley*, for introducing me to the fascinating world of optics research.

My PhD journey has been enriched by the company, friendship, and mentorship of my work colleagues in the Laser Analytics Group. Specifically, I would like to thank:

*Olli* and *Che*, for countless adventures, Garden Restaurant lunches, coffee clubs, and thesis support zooms,

*Dr. Romain Laine*, for his infinite patience when guiding me through the writing of my first paper, and for teaching me what rigour in science looks like,

*Dr. Amberley Stephens*, for her boundless kindness and patience for editing my work,

*Dr. Bo Meng* and *Dr. Katharina Scherer*, for the long hours at the microscope ("turning the wheel"), and for their kindness, mentorship, and camaraderie,

My former colleagues *Craig*, *Marcus*, *James*, *Chris*, *Florian*, *Colin*, and *Nathan*, for welcoming me into the group and introducing me to British culture,

*Craig Russell*, for showing me a fish,

*Miranda*, *Eric Bringley*, *Lucia*, *Sara*, *Roxine*, and *Ezra* for their friendship.

The *Russell Court* family for all the dinners, adventures, cycling trips, parties, and above all, making Cambridge feel like home the past four years.

Finally, I want to thank my parents, *Roberto* and *Maria Consuelo*, and my siblings *Mariana* and *Roberto* for their unconditional love and support through my studies. This thesis is dedicated to them, and to the rest of the Vallejo and Ramirez family, on whose shoulders I continue to stand.

# Table of contents

<b>Abstract</b>	<b>vii</b>
<b>List of figures</b>	<b>xv</b>
<b>List of tables</b>	<b>xix</b>
<b>1 Introduction</b>	<b>1</b>
<b>I Imaging at the nanoscale</b>	<b>3</b>
<b>2 Tools for studying protein colocalisation at the nanoscale</b>	<b>5</b>
2.1 Motivation . . . . .	5
2.2 Background to single-molecule imaging . . . . .	6
2.2.1 Early single-molecule imaging experiments . . . . .	6
2.2.2 Isolating single molecules inside a diffraction-limited spot . . . . .	8
2.2.3 Single-molecule localisation microscopy . . . . .	9
2.2.4 Single particle tracking . . . . .	19
2.3 Components of image analysis workflows . . . . .	21
2.3.1 Image processing . . . . .	21
2.3.2 Colocalisation analysis . . . . .	28
2.4 Summary . . . . .	36
References . . . . .	37
<b>3 A single-molecule method to study the role of alpha-synuclein at the synapse</b>	<b>49</b>
3.1 Introduction to alpha-synuclein . . . . .	50
3.1.1 Structure and function . . . . .	50
3.1.2 A-syn and lipid vesicles . . . . .	51
3.1.3 A-syn and membrane trafficking . . . . .	51

3.1.4	The role of a-syn at the presynapse . . . . .	55
3.2	Experimental Methods . . . . .	56
3.2.1	Synaptosome preparation and stimulation . . . . .	56
3.2.2	Labelling to visualise synaptic proteins and vesicle uptake . . . . .	56
3.2.3	Multi-colour <i>d</i> STORM to study synaptic proteins in synaptosomes .	58
3.2.4	Statistical analysis . . . . .	63
3.3	Synaptosomes as a model to study vesicle recycling . . . . .	63
3.4	Development of SynaptoAnalysis . . . . .	65
3.4.1	Registration of multi-colour single-molecule data . . . . .	66
3.4.2	Automatic detection of synaptosomes . . . . .	69
3.4.3	Colocalisation analysis . . . . .	75
3.4.4	Cluster analysis using the RMSD from centroid . . . . .	77
3.5	Computer model of a synaptosome for generating synthetic SMLM validation data . . . . .	83
3.5.1	Computer model of a synaptosome . . . . .	84
3.5.2	Simulating extreme cases of spatial co-occurrence in synthetic synaptosome data . . . . .	85
3.5.3	Measuring cluster size in synthetic synaptosome data using Ripley's H-function and the RMSD metric . . . . .	88
3.6	Experimental results . . . . .	90
3.6.1	A-syn's co-occurrence with internalised vesicles is temperature independent, compared to that of VAMP-2 . . . . .	90
3.6.2	A-syn and VAMP-2 clusters increase in size in calcium-depleted conditions . . . . .	92
3.7	Summary and Discussion . . . . .	93
3.7.1	A-syn and VAMP-2 associate with different vesicle recycling mechanisms . . . . .	95
3.7.2	A-syn and VAMP-2 clusters increase in size in the absence of calcium	97
3.7.3	SynaptoAnalysis is a useful tool for analysing multi-colour SMLM data from synaptosomes . . . . .	97
3.7.4	Synthetic SMLM data from a model synaptosome is useful to benchmark cluster analysis methods . . . . .	99
3.7.5	Limitations of multi-colour <i>d</i> STORM to study protein distributions in synaptosomes . . . . .	100
3.8	Acknowledgment of contributions . . . . .	101
	References . . . . .	102



<b>4</b>	<b>A study of the colocalisation between HIV and the ESCRT-II protein EAP45</b>	<b>115</b>
4.1	Background to HIV and the ESCRT machinery . . . . .	116
4.1.1	ESCRT proteins are used for membrane remodelling . . . . .	116
4.1.2	HIV sequentially recruits ESCRTs during budding . . . . .	117
4.1.3	The role of ESCRT-II in HIV budding is unclear . . . . .	117
4.2	Imaging and analysis methods to study Gag and EAP45 at the plasma membrane	118
4.2.1	dSTORM imaging . . . . .	118
4.2.2	Validation of a SNAP-tag label for EAP45 . . . . .	121
4.2.3	Transfection of GFP and SNAP-tag to label Gag and EAP45 . . . . .	122
4.2.4	Nearest-neighbour distance analysis . . . . .	124
4.2.5	Single particle tracking of Gag and EAP45 particles in HeLa cells . . . . .	126
4.2.6	Co-moving frame analysis . . . . .	127
4.2.7	Statistical analysis . . . . .	128
4.3	Results . . . . .	129
4.3.1	Maximal experimentally available colocalisation . . . . .	129
4.3.2	FL EAP45 and ALIX show similar co-occurrence with Gag . . . . .	130
4.3.3	FL EAP45 shows a higher co-occurrence with Gag than functionally compromised EAP45 mutants . . . . .	131
4.3.4	Single particle tracking of live HeLa cells shows transient co-occurrence of Gag and EAP45 . . . . .	135
4.4	Summary and Discussion . . . . .	136
4.4.1	SNAP-tag and a GFP fusion provide higher specificity than conven- tional antibodies to fluorescently label EAP45 and Gag, respectively	136
4.4.2	The EAP45 N-terminus is necessary for colocalisation with Gag . . . . .	139
4.4.3	Three potential dynamic interactions between EAP45 and Gag were observed . . . . .	141
4.5	Future work . . . . .	141
4.6	Conclusion . . . . .	143
4.7	Acknowledgment of contributions . . . . .	143
	References . . . . .	144

## **II Imaging at the mesoscale** **149**

<b>5</b>	<b>Introduction to Optical Projection Tomography</b>	<b>151</b>
5.1	Volumetric imaging techniques for biological tissue . . . . .	151
5.1.1	Biomedical tomographic imaging techniques . . . . .	152

5.1.2	3D optical imaging techniques . . . . .	152
5.1.3	Optical Projection Tomography . . . . .	157
5.2	Principles of Optical Projection Tomography . . . . .	160
5.2.1	Overview of OPT working principle . . . . .	160
5.3	Advances in OPT . . . . .	171
5.3.1	Acquisition time reduction and compressed sensing . . . . .	171
5.3.2	Resolution improvement . . . . .	172
5.3.3	Applications of OPT . . . . .	173
5.4	Summary . . . . .	173
	References . . . . .	174
<b>6</b>	<b>Development of a low-cost, open-source OPT system to image large organ sam-</b>	
	<b>ples</b>	<b>185</b>
6.1	Motivation . . . . .	186
6.2	Materials and Methods . . . . .	187
6.2.1	Animal perfusion and tissue preparation . . . . .	187
6.2.2	Clearing and mounting whole lungs for imaging . . . . .	188
6.2.3	Experimental set-up . . . . .	189
6.3	Characterisation of the software and hardware components . . . . .	190
6.3.1	OptiJ hardware . . . . .	190
6.3.2	OptiJ software . . . . .	192
6.4	Results . . . . .	195
6.4.1	Visualisation of primary and higher order airways in different mouse lung lobes using eOPT . . . . .	195
6.4.2	Brightness and signal-to-noise ratio comparison between the two labelling strategies . . . . .	197
6.4.3	Histogram matching for volumetric data visualisation . . . . .	199
6.4.4	Quantifying resolution using Fourier Ring Correlation . . . . .	199
6.5	Summary and Discussion . . . . .	203
6.6	Future work . . . . .	206
6.7	Perspectives . . . . .	207
6.8	Acknowledgement of contributions . . . . .	208
	References . . . . .	209
	<b>Conclusions</b>	<b>213</b>
	<b>Publications</b>	<b>217</b>

# List of figures

2.1	Centroid localisation for a point emitter . . . . .	7
2.2	Photo-activatable and photo-switchable probes . . . . .	8
2.3	Diagram showing separation of fluorescent emissions in time . . . . .	10
2.4	Image processing pipeline for SMLM . . . . .	13
2.5	Effect of labelling density, localisation precision, and long linker lengths in SMLM reconstructions . . . . .	16
2.6	Photo-switching of organic dyes in the presence of a reducing agent . . . . .	18
2.7	Diagram with main steps for single particle tracking . . . . .	20
2.8	Chromatic aberrations can bias SMLM colocalisation studies . . . . .	22
2.9	Demonstration of Otsu's method . . . . .	26
2.10	Colocalisation analysis by area overlap . . . . .	30
2.11	Using Ripley's K-function to analyse spatial point patterns . . . . .	34
2.12	Illustration of Voronoi diagrams . . . . .	35
3.1	Diagram of a-syn's structure . . . . .	50
3.2	Curvature affects packing of phospholipid groups . . . . .	51
3.3	A-syn's double anchor mechanism . . . . .	52
3.4	Vesicle recycling mechanisms . . . . .	54
3.5	Synaptosome preparation procedure . . . . .	57
3.6	Microscope used for single-molecule localisation . . . . .	59
3.7	Improvement in pixel uniformity using top-hat beam shaper . . . . .	60
3.8	Schematic of HILO implementation . . . . .	61
3.9	Synaptosome model to visualise internalised vesicles . . . . .	64
3.10	Data processing pipeline using SynptoAnalysis . . . . .	65
3.11	Registration to correct for chromatic aberrations in SMLM . . . . .	68
3.12	Filtering localisation data on fit parameters and point density . . . . .	71
3.13	Segmentation to remove small elements from reconstructed images . . . . .	72
3.14	Percentage of detected synaptosomes as a function of colour overlap . . . . .	73

3.15	Summary of synaptosome detection steps . . . . .	74
3.16	Validation test for colocalisation measurements . . . . .	78
3.17	RMSD simulations with jittered points inside a circle . . . . .	80
3.18	Influence of localisation uncertainty on the RMSD measurement of cluster size	81
3.19	Errors in the RMSD as a function of point jitter and cluster size . . . . .	82
3.20	Computer model for a synaptosome . . . . .	84
3.21	Testing colocalisation in synthetic synaptosome SMLM data . . . . .	87
3.22	Simulated cluster size measurements on synthetic synaptosome data . . . .	89
3.23	Reconstructed images of synaptosomes detected with SynaptoAnalysis . . .	90
3.24	Colocalisation between mCLING/a-syn, and mCLING/VAMP-2 using the WOC . . . . .	91
3.25	Colocalisation between mCLING/a-syn, and mCLING/VAMP-2 using Man- ders' individual coefficients . . . . .	92
3.26	Cluster size measurements for a-syn and VAMP-2 . . . . .	94
4.1	Sequential recruitment of ESCRTs by Gag during HIV budding . . . . .	116
4.2	Measuring the co-occurrence between Gag and EAP45 using <i>d</i> STORM . .	120
4.3	Staining EAP45 and Gag with specific and isotype control antibodies . . . .	121
4.4	Verification of the SNAP EAP45 labelling in cytokinesis . . . . .	122
4.5	Verifying the specificity of the SNAP-tag label for EAP45 . . . . .	124
4.6	Set up for imaging EAP45 and Gag at the plasma membrane using TIRF microscopy . . . . .	125
4.7	Co-moving frame analysis to analyse the temporal interaction between Gag and EAP45 . . . . .	128
4.8	TIRF image of GFP-Gag with an anti-GFP AF647 nanobody . . . . .	130
4.9	Co-occurrence analysis between ALIX/Gag and SNAP-FL EAP45/Gag . .	131
4.10	Functionally compromised mutants of EAP45 and Gag . . . . .	132
4.11	Co-occurrence of Gag and functionally compromised EAP45 mutants in fixed HeLa cells . . . . .	133
4.12	Co-occurrence of Gag and EAP45 in fixed HAP1 EAP45 KO cells . . . . .	134
4.13	SPT data shows transient co-occurrence of EAP45 and Gag . . . . .	137
4.14	Co-moving frame analysis reveals three classes of EAP45 movement relative to Gag . . . . .	138
4.15	Proposed models for the role of ESCRT-II in HIV budding . . . . .	140
5.1	Schematic of a general OPT setup. . . . .	159
5.2	Radon transform operation . . . . .	162

5.3	Illustration of the Fourier Slice Theorem . . . . .	163
5.4	FBP modelled as a filtering operation . . . . .	165
5.5	Back-projection with and without filtering . . . . .	166
5.6	FBP visualised as smearing projections back through the image plane . . .	167
5.7	Hot or dead pixels on the sensor cause ring artefacts in FBP reconstructions	170
5.8	Incorrect axis of rotation estimation causes halo-like artefacts . . . . .	171
6.1	OptiJ workflow . . . . .	186
6.2	Sample embedding and mounting for OPT imaging . . . . .	188
6.3	OptiJ hardware set up . . . . .	190
6.4	OptiJ plugin workflow description . . . . .	193
6.5	Compensation for mechanical jitter in OptiJ . . . . .	194
6.6	eOPT reconstructions of a left lobe labelled for Surfactant Protein C . . . .	196
6.7	eOPT reconstructions of a right medial lobe labelled for TTF-1 . . . . .	198
6.8	Brightness and SNR comparison between the two epithelial cell labelling strategies . . . . .	200
6.9	Histogram matching of cross-sectional slices for volumetric visualisations of OPT lung data . . . . .	201
6.10	Characterisation of OptiJ post-reconstruction resolution using three different methods . . . . .	202
6.11	FRC resolution as a function of the number of projections acquired . . . . .	204



# List of tables

2.1	Summary of the main SMLM techniques . . . . .	11
2.2	Review of software for SMLM colocalisation and cluster analysis . . . . .	36
3.1	Ingredients for photo switching buffer used in <i>d</i> STORM experiments. . . . .	62
3.2	Filtering values for single-molecule localisation parameters. . . . .	70
3.3	Parameters used to simulate SMLM data from an average synaptosome . . . . .	86
5.1	Comparison of volumetric imaging techniques. . . . .	158





# Chapter 1

## Introduction

The objective of my doctoral work was to develop quantitative imaging and analysis tools for visualising phenotypes in models of disease. A large part of this work focused on imaging molecular level events that lead to disease, such as the clustering of a protein involved in Parkinson's Disease, and the role of a protein unit in the budding of the human immunodeficiency virus (HIV). At the macroscopic level, diseases can cause physical phenotypes that affect the appearance of entire organs; the damage that the COVID19 virus does on lung tissue is one such example. The length scales, optical properties of tissue, and sample preparation challenges for mesoscale (1-20 mm) imaging are vastly different, and other techniques are required to address them. In the second part of my work, I describe a method which permits the 3D visualisation of organ-scale phenotypes via tomographic reconstruction.

Because the parts of my thesis work are different in technology and application, I have organised the thesis into two parts: Part I describes the development of single-molecule imaging and analysis methods to examine the colocalisation of membrane proteins in cellular models of Parkinson's Disease and HIV. Part II describes the development of a low-cost, open-source 3D imaging tool to investigate phenotypes at the organism scale. The focus of the thesis is on technology and software developments driven by biological questions, as part of various collaborations. This thesis assumes familiarity with the fundamentals of fluorescence microscopy.

In Chapter 2, the use of fluorescence microscopy to infer biomolecular interaction is briefly introduced, followed by the background to single-molecule imaging techniques. A historical description of how advances in the detection of single molecules and the spreading of fluorescent emissions in time led to single-molecule localisation microscopy (SMLM) is provided. The background to the image processing and analysis tools used to examine colocalisation and clustering of proteins from single-molecule imaging data is then presented,

including a review on colocalisation methods for both conventional and single-molecule imaging techniques.

Chapter 3 describes the use of SMLM to investigate the role of alpha-synuclein (a-syn) in vesicle recycling mechanisms. A-syn and its controversial role at the presynapse are introduced, and synaptosomes are presented as an effective model to image synaptic proteins. A software package developed to automatically detect synaptosomes from SMLM data, and to perform colocalisation and cluster analysis is described in detail. A-syn was found to colocalise with internalised vesicles in a temperature independent manner, suggesting it could be associated with non-canonical vesicle recycling mechanisms. Additionally, both the clusters of related synaptic protein VAMP-2, and those of a-syn were found to increase in size in calcium-starved conditions.

Chapter 4 presents a study on the colocalisation of HIV and the ESCRT-II protein EAP45 during viral egress from the cell. HIV is known to recruit the ESCRT-I and -III complexes during egress, but ESCRT-II was hitherto considered dispensable. Building on biochemical work from collaborators showing the necessity of ESCRT-II for efficient viral budding, an imaging and analysis workflow was developed to quantify the colocalisation between HIV and truncated EAP45 mutants at the plasma membrane. The N-terminus of EAP45 was found to be necessary for colocalisation with the HIV protein Gag. The dynamics of the interaction between Gag and EAP45 were then explored using single particle tracking, and a method to visualise the correlated motion of two moving particles was developed. Three types of interaction were observed between EAP45 and Gag, lasting from 10 seconds up to 6 minutes.

Chapter 5 begins with a review of volumetric bioimaging techniques. Optical Projection Tomography (OPT) is introduced as a bridging method between medical imaging approaches commonly used with organs and large organisms, and 3D light microscopy techniques applied to image tissue slices or cells. OPT can generate cross-sectional 3D reconstructions of transparent, mesoscale objects (~1-20 mm) with micrometre resolution; both anatomical and fluorescently labelled features can be imaged. The working principles of OPT and the reconstruction process are described, along with applications and advances in the field.

Chapter 6 describes the development of a low cost, open source OPT system to image large organ samples, with freely available software plugins to reconstruct the acquired data. The characterisation of the hardware and software components is described, with examples of usage and demonstrations on how to correct common OPT artefacts. The system was applied to image cleared mouse lung samples in a proof-of-concept study of chronic obstructive pulmonary disease (COPD). Two different epithelial cell staining strategies were compared for visualizing high order airways. Detailed instructions for reproducing the set up are provided, to foster a wider adoption of OPT by future researchers.

# **Part I**

## **Imaging at the nanoscale**



# Chapter 2

## Tools for studying protein colocalisation at the nanoscale

### 2.1 Motivation

In biology, interactions amongst different proteins are the basis for most cellular structures and activities [1]. Biologists often investigate whether two or more proteins inside a cell interact, in order to understand pathways of cellular function. Knowledge of these pathways can be used to design and test therapeutics against disease.

Experiments using biochemical assays [2, 3] and mass spectrometry [4] can provide evidence for protein-protein interactions. Different experimental methods are often used in conjunction to provide multiple indications of association between proteins. The spatial proximity between proteins and other sub-cellular structures is considered a hallmark of many types of physical and chemical interactions [5], although spatial proximity does not warrant molecular interaction. Regardless, visual evidence of spatial correlations together with biochemical and physical assays can provide strong evidence for interactions. Fluorescence microscopy is a common method to image two sub-cellular structures of interest and produce visual evidence of their spatial proximity, or colocalisation. In fluorescence microscopy, proteins are tagged with fluorescent dyes. Upon excitation at the appropriate wavelength, these dyes emit photons, which are then recorded by a detector that registers the positions of the emitters in space. The colocalisation between the recorded signal for each fluorescent label can be quantified using image processing and analysis algorithms. These algorithms, often arranged in groups known as workflows [6], are used to determine the likelihood that two or more fluorescently labelled structures are in contact.

Biological questions require tailored microscopy and analysis workflows; there is no one approach to address all questions [6]. The first part of this thesis comprises the development of two novel imaging and analysis toolkits designed to understand how sub-cellular structures colocalise in Parkinson's Disease and HIV *in-vitro* models. In the Parkinson's Disease model, the colocalisation of two synaptic proteins with synaptic vesicles is studied in the context of vesicle endocytosis (membrane uptake), whereas in the HIV model the colocalisation between a cytosolic protein, EAP45, and the main HIV protein, Gag, is studied in the context of viral egress from the cell (membrane budding and release).

This introductory chapter provides the background to the imaging and analysis techniques used in both colocalisation studies. First, single-molecule localisation microscopy is introduced as a method to study colocalisation at nanometre-scale resolution in fixed cells, followed by single particle tracking to study the dynamic interactions between objects in live cells. The image processing and analysis methods used in both studies are then introduced, including the correction of chromatic aberrations resulting from using multiple fluorescent labels with different emission spectra, segmentation to separate regions of interest from background and noise, and colocalisation analysis.

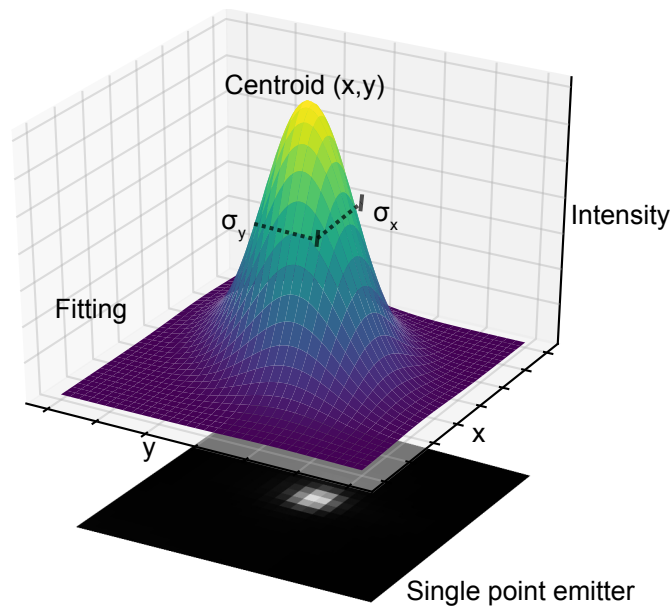
## 2.2 Background to single-molecule imaging

### 2.2.1 Early single-molecule imaging experiments

In the early 1990s, thanks to advances in labelling methods for biological structures, sensitivity of imaging detectors, and brightness of fluorescent probes, single-molecule imaging was first attained with the detection of single emitters in crystals at cryogenic temperatures [7, 8]. Experiments detecting single fluorescent molecules [9] and proteins at room temperature [10, 11] closely followed. These early studies leveraged the knowledge that the centre position of a single fluorophore can be localised much more precisely than the diffraction-limited image of the fluorophore captured by a pixelated detector. If a single emitter is considered a point source, then its blurred image on a detector (Fig. 2.1) corresponds to the point spread function (PSF) of the imaging system. The detected PSF can be fitted to a model function, such as an Airy Disk pattern (a Bessel function of the first kind) or a 2D Gaussian of the form:

$$f(x,y) = A \exp \left( - \left( \frac{(x-x_0)^2}{2\sigma_x^2} + \frac{(y-y_0)^2}{2\sigma_y^2} \right) \right) \quad (2.1)$$

where  $(x_0, y_0)$  are the centre coordinates,  $A$  is the amplitude,  $\sigma_x^2, \sigma_y^2$  are the variances in  $x$  and  $y$ , respectively. The centre of the Gaussian fit provides an estimate for the centre position of the emitter, with a localisation uncertainty estimated by  $\sigma_x$  and  $\sigma_y$ , as demonstrated with the diagram in Fig. 2.1. These centroid-finding methods were developed and first applied in video microscopy, to localise beads attached to motor proteins with nanometre precision [12]. The ability to video track the motion of single particles, coupled with the detection of single molecules at room temperature by Betzig [10] and later Dickson [11], were important milestones in the development of single particle tracking (SPT) techniques, which will be described later in this chapter.



**Fig. 2.1** The centre position of a point emitter can be localised with very high precision by fitting the image of the emitter to a model function, such as a 2D Gaussian.

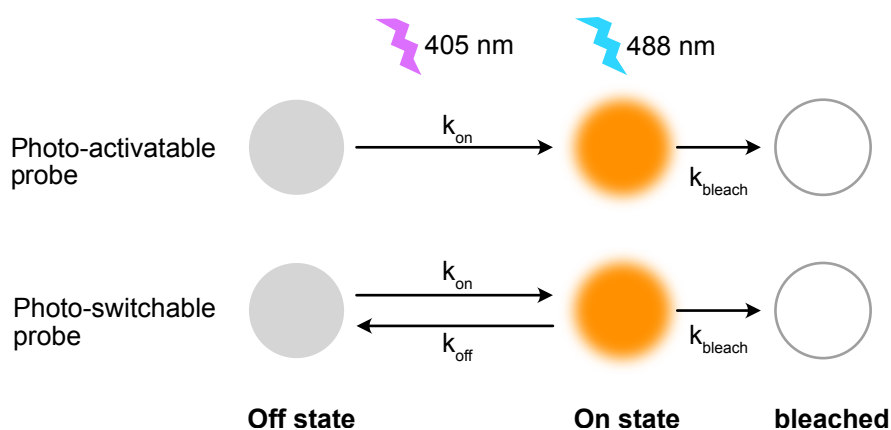
Later in the early 2000s, Thompson et al. [13] derived an equation which describes the limit to how precisely the centroid of a single molecule can be localised. This localisation precision is largely determined by the number of photons collected in each area; the more photons collected, the better the precision of the fit. With negligible noise (shot-noise limited), this localisation precision  $\Delta x$  is roughly given by:.

$$\Delta x \approx \frac{\sigma}{\sqrt{N}} \quad (2.2)$$

where  $\sigma$  is the width of the diffraction-limited spot and  $N$  is the number of photons collected [13].

### 2.2.2 Isolating single molecules inside a diffraction-limited spot

The molecular resolution approach to fluorescence microscopy was further developed by Betzig in 1995, by proposing that individual molecules clustered within the diffraction-limited spot of an imaging instrument could be isolated by one or more defining optical characteristics such as fluorescence spectra, lifetime, or photochromic state, and then localised with a much higher precision [14]. The random fluorescent emission oscillation, or blinking, of fluorescent molecules turned out to be a promising measure to separate the emission from neighbouring molecules in time. Fluorophores had been previously observed to exhibit frequent transitions between different energy states [9, 7, 8], however it was Dickson et al. [11] who first reported the reversible photo-switching of single fluorophores and fluorescent proteins. A diagram of the photo-activation and photo-switching principle of fluorophores is shown in Fig. 2.2.



**Fig. 2.2** The principle of photo-activatable (PA) and photo-switchable (PS) probes. PA probes are activated from a dark *off* state to a bright *on* state by low intensity illumination with a UV laser, and can be subsequently imaged with a readout laser (e.g. a 488 nm laser) until photo-bleached. PS probes reversibly transition between a dark *off* state and a bright *on* state for multiple cycles (blinking) until photo-bleached.

### Centroid localisation applied to single-molecule biophysics

While the development of new and improved photo-activatable and photo-switchable fluorescent probes continued, centroid localisation techniques were being applied to track and image an increasing number of single, sparse molecules. Early applications of centroid localisation for single molecules included the dynamic imaging of single cyanine dyes attached to a molecular motor protein, Myosin V, with nanometre accuracy ( $\sim 1.5$  nm lateral) [15], and single quantum dots (QDs) attached to glycine receptors at the synapse [16]. The QDs were distinguished by their blinking properties, and their centroids were localised to 5-10 nm precision. Following this work, other groups applied centroid localisation to image several (two



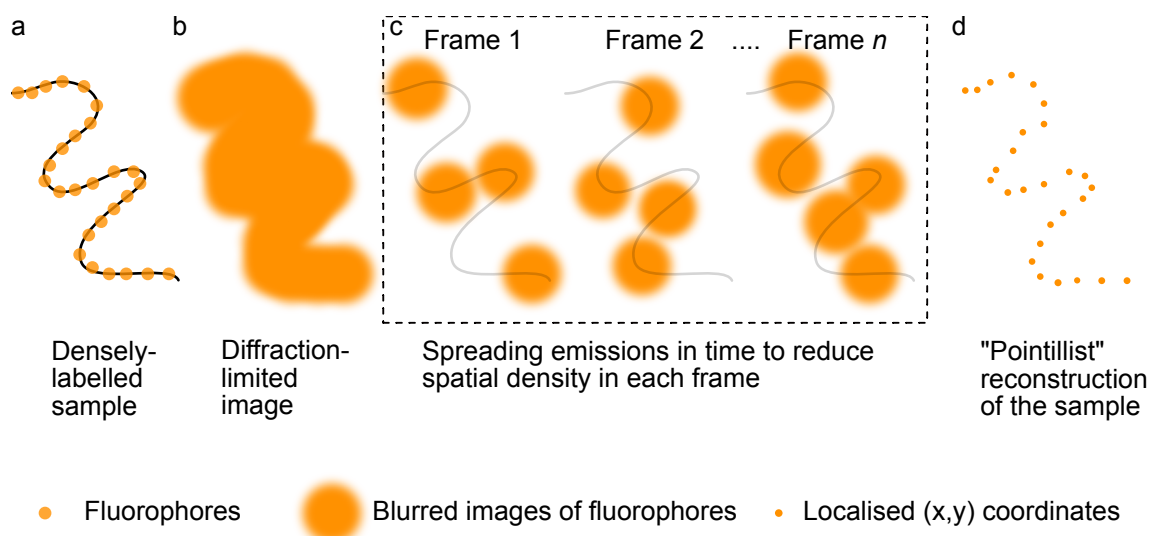
to five) fluorescent dye molecules tethered to DNA scaffolds within a diffraction-limited area. Nanometre localisation accuracy was achieved by sequentially localising and photobleaching each dye on the scaffold [17, 18]. In 2005, Lidke et al. were the first to apply the temporal separation of fluorophore emission to precisely localise the emitter positions in space, and subsequently reconstruct an image consisting of individually localised points [19]. This technique was termed “pointillism”, and it was demonstrated using the blinking properties of QDs to accurately locate single molecules in groups of up to four emitters within 30 nm of each other [19]. None of these methods could yet translate into an increase in traditional imaging resolution, since it was still unfeasible to resolve more than a few emitters within a diffraction-limited region. Nevertheless, having optimised the imaging and sub-diffraction localisation of sparse, isolated single molecules, the next step was to detect single molecules at high molecular density.

### 2.2.3 Single-molecule localisation microscopy

#### Spreading fluorescent emission in time to circumvent the spatial density problem

To resolve biological structures using fluorescence microscopy, dense labelling of a sample’s fine features with fluorophores is often required such that a complete image of the sample structure can be recorded. Lidke et al.’s [19] insight to spread the fluorescent emission from multiple sources in time was essential to circumvent the spatial density problem; by keeping the number of active (*on* state) probes at any point in time sufficiently low, this sparse subset could be imaged as isolated single molecules. A full molecular image of the sample structure could then be reconstructed by imaging a sparse subset of fluorophores in each frame over thousands of frames. Figure 2.3 illustrates the concept of separating emissions in time to address the spatial density problem. To implement this concept, the remaining challenge was to reliably switch fluorophores between the *on* and *off* state.

The photo-switching and photo-activation properties observed in fluorescent proteins and organic dyes during the single-molecule experiments in the 90s and early 2000s were leveraged to control the number of active probes at any point in time. In the case of fluorescent proteins, the characterisation of the photo-switching behaviour of the green fluorescent protein (GFP) [11] led to the optimisation of many fluorescent protein alternatives for better efficiency, longevity, and brightness in the blinking process [20]. In the case of organic dyes, experiments to improve fluorophore photo-stability in single-molecule Förster Resonance Energy Transfer (smFRET) [21, 22] led to the discovery of thiol-induced Cy5 switching off [23]. Cy5 could be subsequently reactivated using a second laser line enhanced by the proximity of a donor probe, such as Cy3 [24], or by ultraviolet (UV)



**Fig. 2.3** Diagram showing how the spatial density problem can be addressed by separating fluorescent emission in time. (a) A structure densely labelled with fluorophores can result in a diffraction-limited image with overlapping fluorophores, such as the blurred blob shown in (b). By spreading the emission of the different fluorophores such that only a sparse subset is on in any given frame (c), a pointillist reconstruction of the sample (d) can be assembled by sequentially localising the centroids of the sparse molecules and superposing them.

excitation. Many other conventional organic dyes have now been demonstrated to photo-switch controllably under suitable buffer chemistry and irradiation [25]. The challenge of localising multiple emitters in a diffraction-limited area was addressed nearly simultaneously in 2006 by three different techniques [26, 27, 28]. These techniques combined active control of fluorophore blinking to separate the signal from individual probes in time with highly precise centroid localisation methods to circumvent the resolution limit in space. This mechanism is now widely applied in a series of SMLM techniques, which are photo-activated localisation microscopy (PALM) [26], fluorescence PALM (fPALM) [27], stochastic optical reconstruction microscopy (STORM) [28], direct STORM (dSTORM) [29, 25, 30], ground-state depletion microscopy followed by individual molecule return (GSDIM) [31], and point accumulation for imaging nanoscale topography (PAINT) [32, 33, 34, 35, 36, 37]. Table 2.1 summarises the main SMLM techniques, their working principle, and some advantages and disadvantages of each.

**Table 2.1** Summary of the main SMLM techniques.

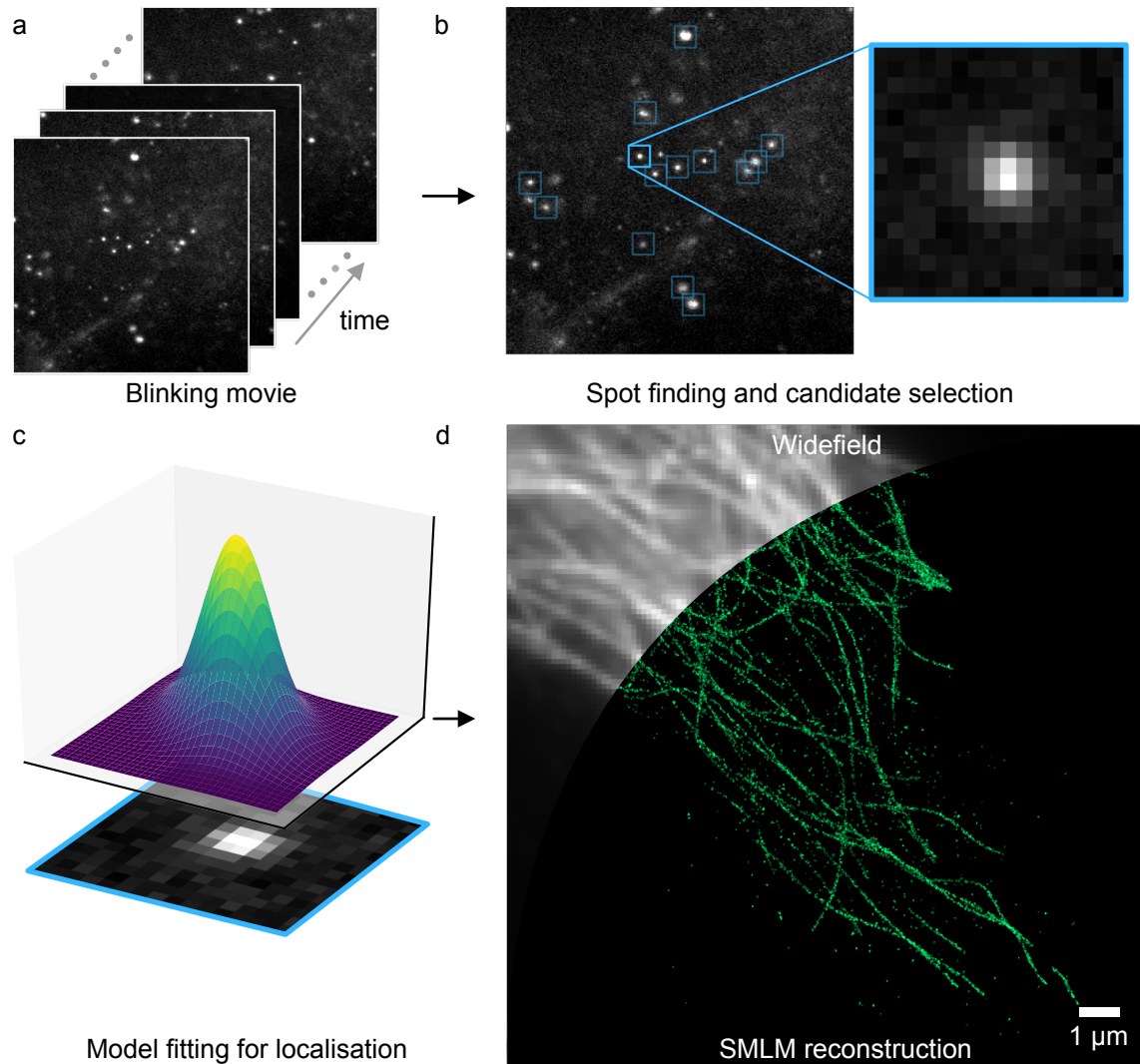
Technique	Working Principle	Advantages	Disadvantages	Reference
PALM	Photo-activatable fluorescent proteins (PA-FPs) are attached to target proteins. PA-FPs, native in the <i>off</i> state, are switched on by an activation laser, and then imaged with a readout laser until most are bleached. This is repeated until the population of inactivated, unbleached PA-FPs is depleted (the related technique fPALM [27] works similarly).	Endogenous labelling with PA-FPs yields high specificity and a negligible linker length between the fluorophore and the target protein.	Requires genetically fusing PA-FPs to target proteins.  PA-FPs can only emit a few hundred photons before they bleach, which decreases the attainable localisation precision.	Betzig et al. 2006 [26].
STORM	Photo-switchable dye pairs, an activator and reporter such as Cy3 and Cy5, are attached to target molecules. The dye pair can be deactivated ( <i>off</i> state) by a far-red laser and activated ( <i>on</i> state) by a green laser. Random sparse subsets of fluorophores can be switched <i>on/off</i> hundreds of times before photobleaching.	Minimises chromatic aberrations for multi-colour acquisitions, since the same reporter dye can be used with multiple different activator dyes.	Requires a dye pair to be conjugated to the target molecule.  Requires timed control of laser pulses to alternate between the activation and readout beam during frame acquisition.	Rust et al. 2006 [28].
dSTORM	Photo-switchable organic dyes are deactivated at the start of the acquisition and stochastically switched between the <i>on/off</i> state with a single laser until photo-bleached. (a similar principle is used for GSDIM [31])	Does not require an activator-reporter dye pair; conventional fluorophores can be used as photo-switchable probes.  Only requires a single laser line for activation and readout	Suffers from chromatic aberrations when using multiple fluorophores.  Conventional antibodies for labelling molecules with organic dyes have long linker lengths, introducing biases in the localisation of single molecules.	Heilemann et al. 2008 [25].
DNA-PAINT	Single stranded DNA oligomers are attached to the target molecule via conventional antibodies (docking strand). Complementary DNA strands conjugated to an organic dye diffuse freely in solution and transiently bind to the docking strand, causing a blinking effect in the detected frames. Implementation of the PAINT principle from Sharonov et al. 2006 [37].	Does not depend on dye photo-physics for achieving temporal separation of emission.  Enables multi-colour imaging for > five channels using different imager strands.	Acquisition process is ~10x slower than dSTORM.  Multi-colour imaging requires repeated imaging + washing cycles that can introduce drift/motion artefacts between colour channels.	Jungmann et al. 2010 [34].

The names of the techniques denote the specific constraint mechanism used to keep the active emitter concentration sparse, and thus avert emitter overlap in any one frame (e.g. photo-activation, photo-switching using intrinsic dark states, or transient binding from collisions in solution). Aside from SMLM, there are other fluorescent super-resolution microscopy techniques which also circumvent the diffraction limit, such as stimulated emission depletion (STED) microscopy and structured illumination microscopy (SIM) [38]. STED microscopy can achieve 50 nm resolution by depleting the fluorescence of emitters around a diffraction-limited region, thereby confining the effective fluorescent emission volume to a sub-diffractive spot [39]. SIM can achieve up to 100 nm lateral resolution by acquiring multiple frames containing spatial frequency information outside the passband of the optical system, and combining them in Fourier space to achieve double the resolution of the raw images. In this thesis, SMLM was chosen specifically because it offers the largest resolution improvement of all super-resolution techniques, and the SMLM microscope available in our laboratory offered the necessary laser excitation lines (488, 568, and 647 nm) for performing multi-colour imaging. For further background on SMLM, Sauer et al. [40] and Deschout et al. [41] provide comprehensive reviews. Li and Vaughan [42] provide an excellent review on switchable probes for SMLM. The following section describes the general working principles for most SMLM techniques, encapsulating the building blocks presented above.

### Working principles of SMLM

Unlike in traditional fluorescence microscopy, in which the fluorescent emission of all fluorophores in a sample is collected simultaneously or by raster scanning across the sample, in SMLM, images are assembled by collecting numerous frames containing sparse, single emitters, and precisely determining the centre position of each emitter over multiple *on/off* switching cycles. A super-resolution image is then generated as a computational reconstruction of the localised fluorophore density. This fundamental concept is illustrated in Fig. 2.4. The temporal separation of fluorescent emission is achieved by using probes which exhibit photo-switching properties, or reversibly bind to target structures.

In practice, SMLM is achieved by first acquiring several thousands of frames of the blinking fluorophores (Fig. 2.4 a), and then processing frames with a localisation algorithm to determine the centre coordinates of each emitter with high precision. The algorithm first detects local maxima in each frame to get approximate localisations of bright emitters (*on* state) [43] (Fig. 2.4 b). These approximate localisation candidates are extracted, and the sub-diffractive localisation of the emitter centroid is accomplished by fitting each candidate's 2D photon distribution to a suitable PSF model, e.g. a 2D Gaussian (Eq. 2.1, Fig. 2.4 c), using



**Fig. 2.4** Image processing pipeline for SMLM. (a) Images of sparse fluorescent emitters are segmented to detect bright spots (b), and each bright spot is then fitted to a model function such as a 2D Gaussian to locate its centre position (c). The localisation density map is then plotted to reconstruct a super-resolution image, shown in (d) with an overlay of the diffraction-limited equivalent image. HFF cells stained for microtubules with Alexa Fluor 647 are shown in this figure.

non-linear least-squares methods or maximum-likelihood estimation [44]. These elements of image processing for localisation microscopy are summarised in Fig. 2.4. Although Eq. 2.2 provides a good estimate in cases with negligible noise, the uncertainty in the localisation precision  $\sigma_x$  is generally calculated using Mortensen's correction [44] of Thompson et al.'s original formula [13], which accounts for dipole (anisotropic) photon emissions and properties of both CCD and EMCCD detectors, written below:

$$\sigma_x^2 = \frac{\sigma_{psf}^2 + \frac{a^2}{12}}{N} \left( \frac{16}{9} + \frac{8\pi b^2 (\sigma_{psf}^2 + \frac{a^2}{12})}{Na^2} \right) \quad (2.3)$$

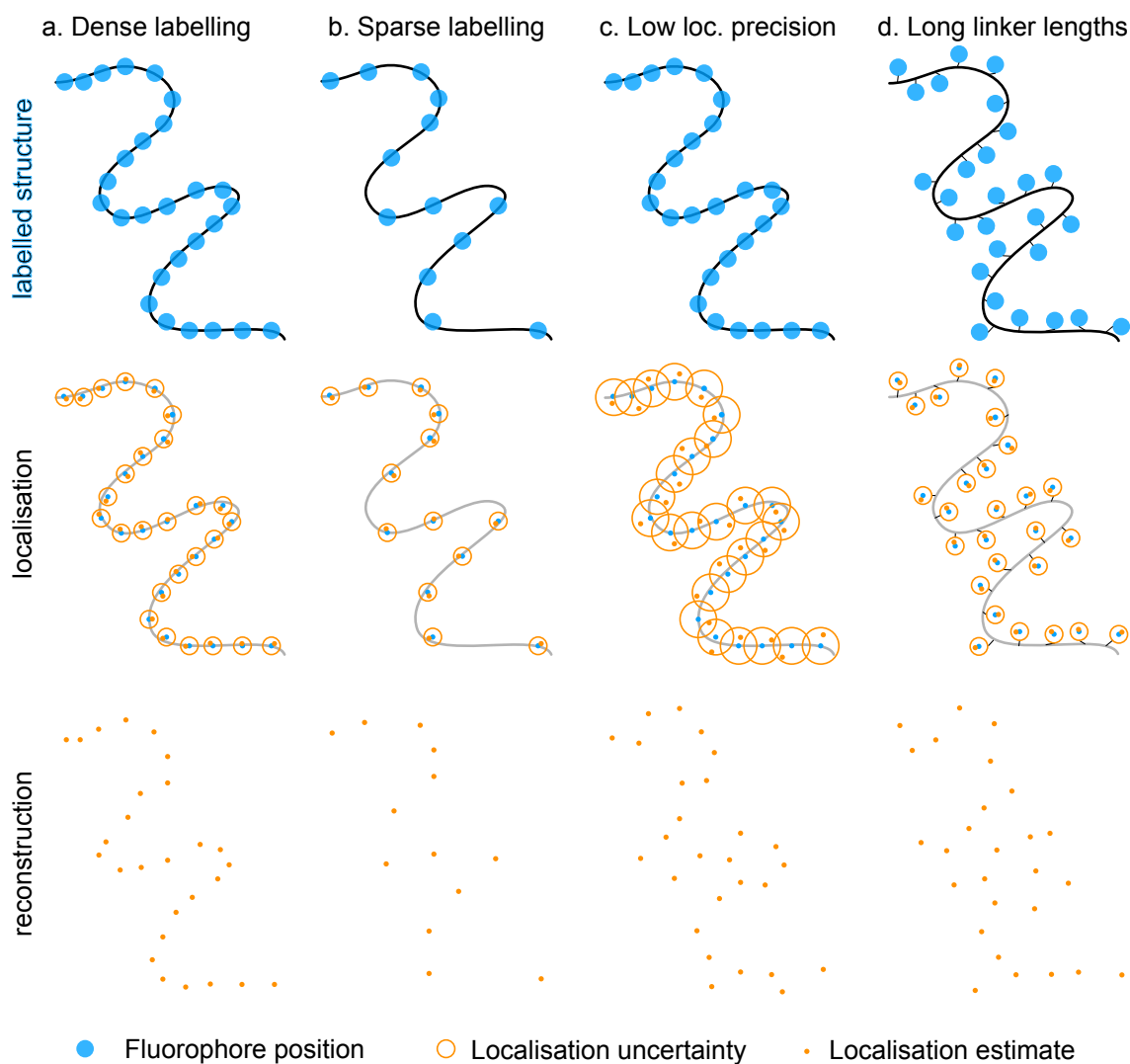
where  $\sigma_{psf}^2$  is the variance of the fitted PSF,  $N$  is the number of photons detected from the emitter,  $a^2$  is the pixel area, and  $b^2$  represents the assumed background photons per pixel. Thompson et al. [13] reported a 30% excess error in localisation precision when using Gaussian masks to fit experimental images of fluorescent beads, compared to the expected localisation value from computer simulations of Gaussian-distributed photons. The factor 16/9 in Eq. 2.3 was empirically determined to compensate for this excess error ( $(16/9)^2 = 133\%$ ). Additionally, Eq.2.3 also treats the assumed background  $b^2$  per pixel exactly by including its product with the pixel area ( $a^2$ ) term, unlike Thompson et al.'s formula which neglects this correction factor in the background term. The result of this procedure is a table containing the  $(x,y)$  coordinates of each emitter, the integrated signal density (in units of photons), the width of the PSF fit, and the localisation uncertainty  $\sigma_x$ . A variety of reconstruction algorithms have been developed, and multiple software packages have been produced and are readily available [45, 46, 47]. A super-resolved image can be then reconstructed by generating a 2D histogram from the localisation table with a pixel size in the range of the localisation precision. Alternatively, each localisation can be plotted as a single point convolved with a Gaussian blur function whose standard deviation is equal to the localisation precision [40]. This concludes the general acquisition and reconstruction process for SMLM, after which image or spatial point pattern analysis techniques can be used to extract further information from the data.

### Image resolution in SMLM

The resolution in an SMLM data set depends on both the localisation precision achieved, the labelling density, and the kinetics of the dye photoswitching [41]. The localisation precision calculated from Mortensen et al.'s formula [44] provides a measure of how well the signals from two neighbouring emitters can be distinguished, similar to the Abbe limit for diffraction-

limited imaging. SMLM techniques regularly achieve localisation precision on the order of ~5-10 nm assuming 1000 detected photons per switching cycle and moderate background noise [40]. At this length scale, the fluorescent probe distribution may significantly differ from the target structure [48]. The target structure must be properly sampled by fluorophores such that it can be faithfully reproduced in the reconstructed image, as demonstrated with the illustrations in Fig. 2.5. The labelling density must be determined depending on the sample structure. For samples with continuous boundaries, e.g. microtubules or mitochondrial membranes, the Nyquist criterion [49] can be applied such that the distance between neighbouring probes must be smaller than half the size of the smallest feature that can be resolved. For samples without continuous boundaries, such as proteins attached to lipid membranes which are only expressed in low copy numbers, or cytoplasmic proteins that are free floating, the labelling density dictated by the Nyquist criterion might not be a useful metric. Regarding photoswitching kinetics, the photoswitching ratio  $r = k_{off}/k_{on}$ , which is ratio between the *off* and *on* states, needs to be high, as this ensures only a small number of fluorophores will be in the *on* state at any given frame, ensuring only single fluorophores will be localised.

Other, more complex metrics of SMLM resolution have been reported such as the information transfer function and Fourier Ring Correlation (FRC). The information transfer function is derived using Fisher information theory [50], and it provides an estimate of how much information a parameter X (e.g. the emitter intensity) can provide about an unknown parameter Y (e.g. the emitter centre position) [51]. This theoretical framework requires *a priori* knowledge of the sample structure and thus may be impractical for some applications [41]. FRC is based on evaluating the spatial frequency correlation between two independently recorded images of the same region and reporting a cut-off spatial frequency based on the correlation decay. FRC has been used for decades in electron microscopy to estimate image resolution, and it was recently adapted for optical nanoscopy [52, 53]. In FRC applied to SMLM data, localisation data tables are divided in two halves, and the correlation between the Fourier transform of each half's reconstruction is calculated over a series of concentric rings in Fourier space. This results in a graph of decaying correlation as a function of spatial frequency. The image resolution can then be obtained as the inverse of the spatial frequency for which the FRC curve drops below a pre-set threshold value. The FRC value should not be used indiscriminately, as it is not an absolute measure of resolution; rather it should be reported along with the localisation precision and the estimated probe density (nearest-neighbour distance) to give an overall estimate of the reconstructed image resolution. Comprehensive reviews are provided by Chao et al. [51] on Fisher information theory, and by Nieuwenhuizen et al. [53] on using FRC for resolution estimation in SMLM.



**Fig. 2.5** Drawings showing the effect of labelling density, localisation precision, and long linker lengths in SMLM reconstructions. Column (a) shows a densely labelled structure with fluorophores localised with high precision yields a faithful reconstruction of the sample. Reconstruction quality and resolution deteriorates with sparse labelling (b), low localisation precision during reconstruction (c), and long linker lengths between the target molecules and the fluorophores used (d). Figure inspired by Deschout et al. [41].



The historical background and working principles of SMLM have now been described. This thesis specifically applied *d*STORM to examine nanoscale protein distributions in Chapters 3 and 4, therefore a brief explanation of the specific constraint mechanisms to temporally separate fluorescent emissions in *d*STORM is included in the following section.

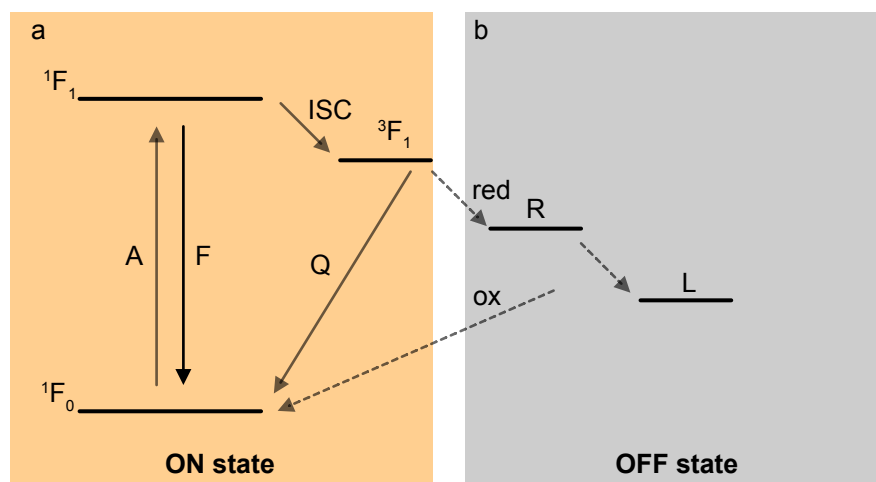
### ***d*STORM working principle**

In *d*STORM, photo-switchable organic dyes are used to achieve temporal separation of fluorescent emission. Target molecules are labelled, and all fluorophores are transferred to a reversible *off* state at the start of the experiment by irradiating the sample with high laser intensity at the peak of the fluorophore's excitation spectrum. A sparse subset of fluorophores will return to the *on* state either stochastically, or after irradiation with a UV source. If the reactivation probability is sufficiently low, then it is highly unlikely that any two fluorophores in the *on* state will be closer together than the diffraction limit. In other words, the lifetime of the *off* state must be substantially longer than that of the *on* state to keep the active emitter concentration sparse. This cycle of activation, localisation, and deactivation is repeated hundreds of times until the dyes are photo-bleached.

Altogether, unlike in STORM in which an activator/reporter dye pair is necessary to switch between *on/off* states (Table 2.1), in *d*STORM single organic dyes are reversibly photo-switched making use of:

1. High laser irradiation to achieve an inter-system crossing between energy states [30].
2. A reducing agent and an oxygen scavenger in buffer solution to control the return rate from the *off* to the *on* state, and ensure a long *off* state lifetime [30].

***d*STORM photo-switching mechanism** Upon illumination at its absorption maxima, an organic dye will be cycled multiple times between its ground state ( $^1F_0$ , Fig. 2.6) and its excited singlet state ( $^1F_1$ , Fig. 2.6), emitting fluorescence photons with every decay from the singlet to the ground state with a lifetime of a few nanoseconds. The dye can then undergo an inter-system crossing (ISC, Fig. 2.6) either stochastically [54] or induced by high irradiation [30] into a triplet state ( $^3F_1$ , Fig. 2.6). From the triplet state, the dye can either be quenched (Q, Fig. 2.6) by molecular oxygen back to the ground state, or further reduced by a thiol such as mercaptoethylamine (MEA) to generate a stable *off* state of up to a few seconds [40]. Both of these reactions are necessary. The occasional quenching of the triplet state by molecular oxygen prevents the dye from being reduced into a dark state every time it reaches the triplet state, thereby allowing for repeated fluorescent emission cycles. Hundreds



**Fig. 2.6** Photo-switching of organic dyes in the presence of a reducing agent. (a) Upon fluorescent excitation, a dye can either cycle between its singlet ground  $^1F_0$  and excited  $^1F_1$  state, emitting fluorescent photons, or it can undergo an intersystem crossing (ISC) to populate the triplet state  $^3F_1$ . (b) The triplet state can either be quenched (Q) by molecular oxygen back to the ground state or reduced (red) by a thiol to produce the fluorophore radical (R). Some dyes can accept a second electron to form the fully reduced leuco form of the dye (L). Both the reduced (R) and leuco (L) forms can return to the ground state by collisions with molecular oxygen, i.e. oxidation (ox).

to several thousands of photons can then be detected prior to reduction of the triplet state by a thiol into a long-lived *off* state.

The reduction of the triplet state can result in either a fluorophore radical (R, Fig. 2.6) or a fully reduced leuco form of the dye (L, Fig. 2.6), which are oxidised to repopulate the singlet ground state  $^1F_0$ . For most rhodamine derivatives (e.g. Atto 647N, Alexa Fluor 568, etc.) the formation of a fluorophore radical R is accompanied by an additional blue-shifted absorption band for the fluorophore around 400 nm. This allows for a photo-induced recovery of the singlet ground state by irradiation with a UV (~405 nm) laser. In the absence of a reducing thiol, the lifetime of the dark triplet state is only a few microseconds [30], and is therefore impractical for SMLM.

In summary, the photo-switching of organic dyes depends on high irradiation to drive inter-system crossings, and a balance of reduction-oxidation (redox) reactions to control the lifetime of the *off* state. The redox reactions depend on the thiol concentration to reduce the triplet state and enable a long-lived *off* state, and on the molecular oxygen concentration not only to quench the triplet state and ensure high photon yield, but also to repopulate the singlet ground state by oxidising the reduced form of the dye. The photo-switching rate also depends on the buffer solution, the use of a UV laser to repopulate the singlet state, and the type of dye, e.g. rhodamine, oxazine, or carbocyanine. Different dyes have distinct

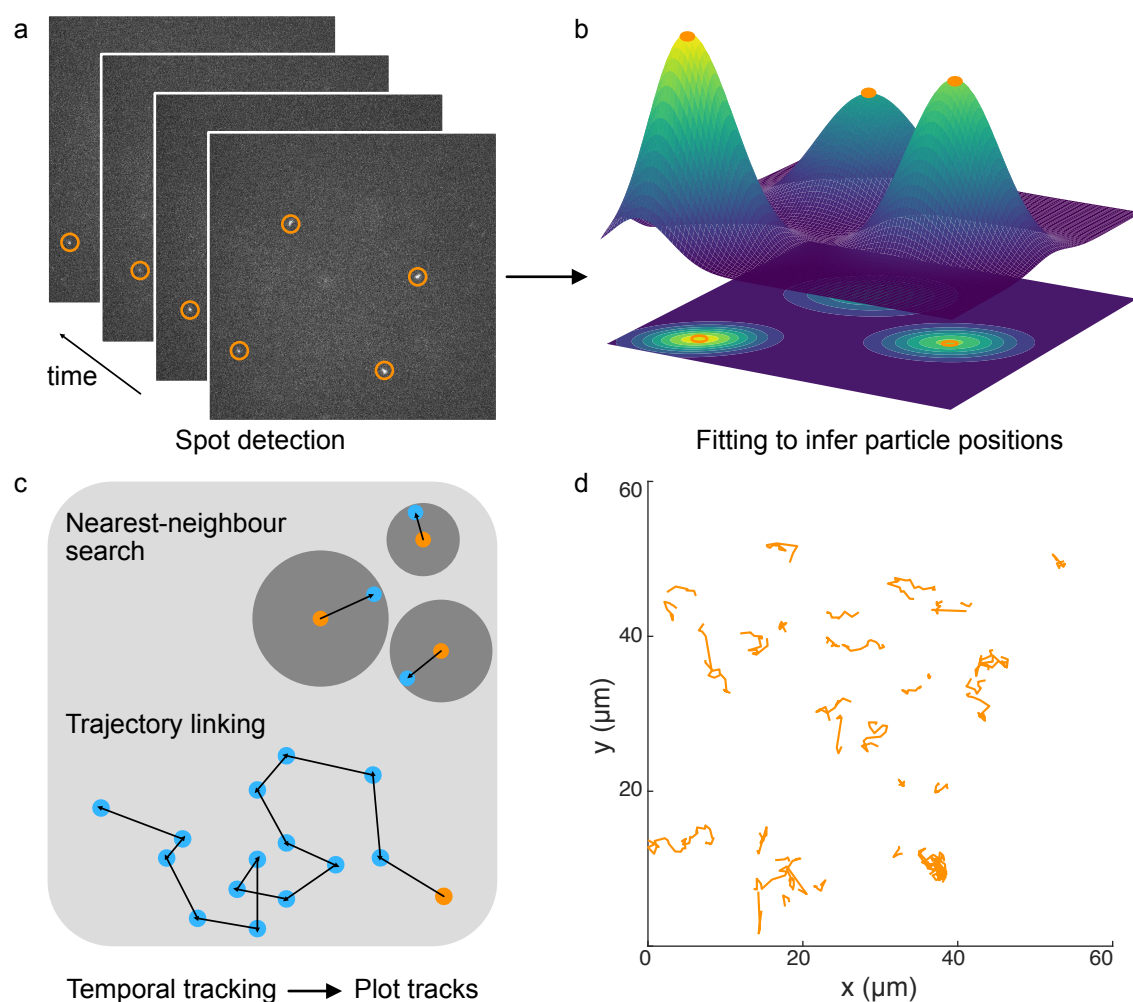
redox properties, and therefore require tailored thiol and oxygen scavenger concentrations to achieve stable, long-lived *off* states and high photon yield. Nevertheless, Dempsey et al. [55] have previously reported that most dyes perform best using both a glucose oxidase and catalase oxygen scavenger, and a thiol such as MEA.

A thorough explanation of the photo-switching chemistry behind *d*STORM can be found in [56, 30, 40, 54]. Additionally, Dempsey et al. [55] report a methodical characterisation of over 20 popular fluorophores for localisation microscopy, and present guidelines for choosing photo-switchable dyes depending on the application. van de Linde et al. [30] provide a meticulous protocol for *d*STORM imaging, with discussions on how to adjust the ratio of fluorophores in the *on* and *off* states, as well as best practices for labelling samples and acquiring *d*STORM data sets.

## 2.2.4 Single particle tracking

The centroid-finding algorithms developed for video microscopy [12], coupled with the detection of single molecules at room temperature [10, 11] as described in the previous section were fundamental to the development of SPT techniques over the past three decades. SPT comprises a family of methods used to record and analyse the dynamics of particles either in isolation or in ensembles. Particle tracking is an important tool for biology, enabling discoveries such as the observation of the motion of motor proteins along the cytoskeleton [15], or the entry, transport, and egress of viruses in cells [57], among many others [58]. SPT can be generally divided into two stages: spatial detection (Fig. 2.7 a-b), and temporal tracking (Fig. 2.7 c-d), both of which share common elements with localisation microscopy.

Spatial detection involves identifying single, sparse molecules at high resolution, and localising their centres with sub-pixel accuracy over multiple frames by fitting the image of each emitter to a model function, e.g. a 2D Gaussian. This is analogous to the spot finding and candidate selection process described for SMLM in Fig. 2.4 b-c, assuming that point-like particles with symmetric images are being observed. In SPT however, the particles observed are dynamic, and therefore localised emitters over each frame will trace displacement trajectories that can be isolated and analysed. The temporal tracking step involves linking detected particles in neighbouring frames to trace their displacement trajectories. The simplest method for particle linking is the nearest-neighbour, which assigns a link to the nearest detected particle in the following frame, as illustrated in Fig. 2.4 c. More sophisticated, multi-frame association methods include multi-hypothesis tracking [60], which is approximated by Jaqaman's popular linear assignment problem (LAP) tracker [61], and combinatorial approaches such as that proposed by Sbalzarini et al. [62]. To illustrate a basic SPT experiment, Fig. 2.7 shows the process of tracking Cy5-tagged mRNA molecules



**Fig. 2.7** Schematic diagram showing the two main stages of SPT: particle detection (a-b), and temporal tracking (c-d). (a) Pixels with local brightness maxima in a time-series stack are identified as candidates, and subsequently fitted with a model function to estimate their centroid (b). (c) The centre coordinates corresponding to the same particle at different times are linked using methods such as the nearest-neighbour. (d) The resulting displacement trajectories, or tracks, can be plotted. The ImageJ plugin TrackMate [59] was used to detect the particles and generate the tracks in this demonstrative experiment.

diffusing in water, imaged using a 1.49NA oil immersion objective. Individual molecules diffusing in and out of the focal plane were detected and fitted (Fig. 2.7 a-b), and their trajectories were linked using Jaqaman's LAP tracker. The resulting tracks were plotted in Fig. 2.7 d, as an illustration of typical random walks by diffusing particles in solution.

The resulting displacement trajectories can then be used to study physical properties of the sample, such as diffusion, transient interactions, trapping, and velocity in time and space. Extensive reviews on SPT can be consulted for further details on specific methods and applications [58, 63]. In this thesis, SPT was specifically applied in Chapter 4 to measure and analyse the transient interaction between HIV and a cytosolic protein, EAP45, during viral egress from the cell.

## 2.3 Components of image analysis workflows

The single-molecule imaging methods described above provide the necessary resolution and sensitivity to examine protein distributions at the nanoscale, both in fixed and dynamic samples. Measuring the colocalisation between different protein distributions, however, requires specialised image analysis methods. This section provides the background to the image processing and analysis methods applied in Part 1 of this thesis.

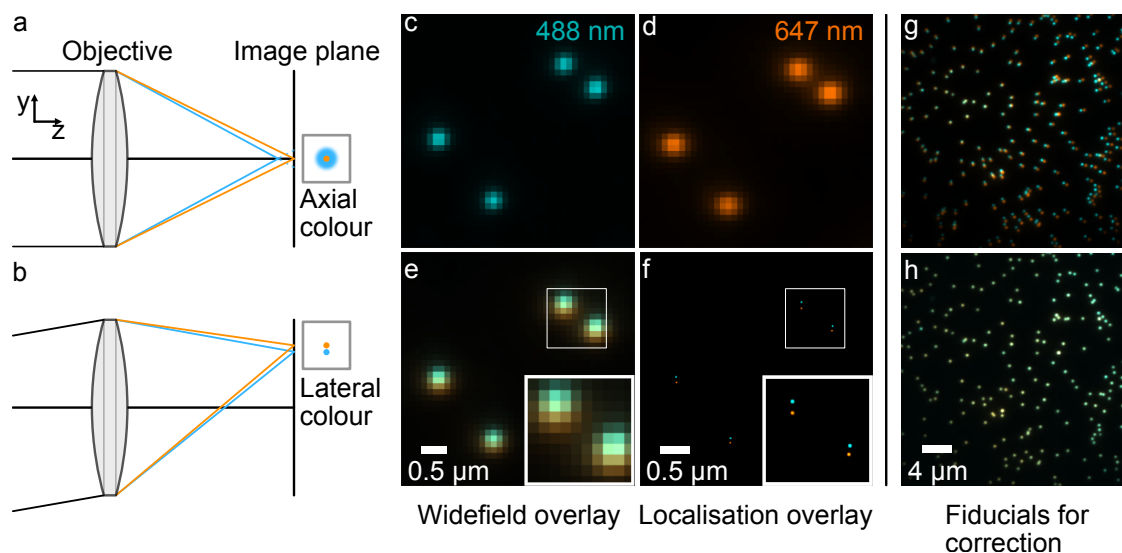
### 2.3.1 Image processing

This subsection introduces image processing methods to correct for chromatic shifts in the acquisition process (registration), and discern between regions of interest and background or noise (segmentation).

#### Chromatic registration

Chromatic aberrations are wavelength-dependent errors in light propagation often caused by the variation in refractive index with wavelength of refractive optical elements [64]. For ray bundles of different wavelengths traveling in an optical system, axial chromatic aberration, or axial colour, is the variation in focal position along the optical axis with wavelength (Fig. 2.8 a), whereas lateral chromatic aberration, or lateral colour, is the change in magnification as a function of field angle for each wavelength (Fig. 2.8 b) [64]. In microscopy, commonly used plan apochromat objectives are designed to correct for spherical aberration, coma, field curvature, and chromatic aberrations for up to five wavelengths [64]. In practice, small residual aberrations are often present even in well-corrected microscope systems due to tilt

of the coverslip, misalignment in the imaging optics between the objective and the detector, and/or refractive index mismatch between the sample media and the objective [65].



**Fig. 2.8** Chromatic aberrations can bias SMLM colocalisation studies. (a) Ray diagram showing axial (a) and lateral (b) chromatic aberrations and their effect on the image of a point emitter at infinity. (c-d) 100 nm beads imaged in two colours to demonstrate lateral chromatic offset. A region was extracted at the edge of the field of view to highlight how, although the beads overlap in the diffraction-limited overlay (e), the localised centres of the cyan and orange emitters may actually appear more than 100 nm away from each other (f). (g) Representative field of view containing sub-diffractive beads as fiducials, which can be corrected to yield the image in (h).

In the context of SMLM, chromatic aberrations are especially problematic, as objective lenses are designed to have “diffraction-limited” performance, which means aberrations are minimised for imaging spatial frequencies the size of, or larger than the diffraction limit. For SMLM systems in which imaging resolution of 20-30 nm is achieved, even the best objective lenses will exhibit pronounced chromatic aberrations for multi-colour imaging [66, 67]. These chromatic offsets can introduce biases or errors in colocalisation studies, therefore they must be corrected or minimised prior to analysing multi-colour SMLM data.

Lateral colour, which has the most prominent effect on multi-colour localisation data for 2D imaging, is illustrated using 100 nm multi-colour beads in Fig. 2.8 c-f. The images were captured using 488 and 647 nm lasers, and an Olympus UAPON 100x 1.49 NA objective, which is corrected for five colours at diffraction-limited performance. Assuming the optical axis is in the centre of the field of view, lateral colour will cause the image of point emitters away from the optical axis to deviate from its true location (Fig. 2.8 b,e,f). For widefield, diffraction-limited imaging, even at the edge of the field of view (Fig. 2.8 e), the images of

the emitters still overlap, however the localised centres of the emitters are distinctly separated by more than 100 nm (Fig. 2.8 f).

Chromatic offset is generally corrected by imaging fiducials, e.g. 50 to 100 sub-diffractive multi-colour fluorescent beads as shown in Fig. 2.8 g, and subsequently generating a mapping function between a chosen reference colour channel and the warped (offset) channel to register the two-colour images, resulting in an image with no chromatic offset as demonstrated in Fig. 2.8 h. This mapping function can be a geometric transformation which compensates for aberrations, e.g. distortions, shifts, or shears, between the positions of the reference and warped fiducial images. For SMLM data sets, the figure of merit for the mapping function is the registration error between the reference and the warped channels. In practice, a registration algorithm is effective when the registration error is smaller than the localisation precision of the data. Registration algorithms have been previously proposed for SMLM by Annibale et al. [66] and Erdélyi et al. [67], which inspired the algorithms developed in Chapter 3 of this thesis to register multi-colour SMLM data sets. The computational intensity of this procedure is relatively low, taking at most a few minutes to register localisation data sets.

### Segmentation

Segmentation partitions images into labelled sub-regions that share similar characteristics. In microscopy, segmentation discriminates between useful regions in a cell (e.g. organelles, membranes, or distinct protein clusters) and background signal. In Chapters 3 & 4, segmentation was used to detect protein distributions of interest and remove unwanted signals that could bias colocalisation analysis. The segmentation operations used in this thesis involved two key steps:

1. Thresholding an image to discern between specific signals and background.
2. Applying morphological operations to remove unwanted regions in the thresholded image.

**Thresholding** Thresholding involves partitioning images into regions based on intensity values. For an image  $f(x,y)$ , pixels above a threshold value  $T$  are assigned a value of 1 (white), and those below a value of 0 (black), as in the equation below:

$$f(x,y) = \begin{cases} 1, & \text{if } f(x,y) \geq T \\ 0, & \text{if } f(x,y) < T \end{cases} \quad (2.4)$$

If  $T$  remains fixed for the entire image, it is a global threshold, whereas if  $T$  changes depending on the pixel region, it is a local threshold. Thresholding can also be understood as a classification problem in which the objective is to minimise the average error incurred in assigning pixels to two or more classes. In this thesis, two different solutions to the classification problem were used for segmentation: Otsu's method [68] in Chapter 3, and Weka segmentation [69] in Chapter 4.

**Otsu's method** Otsu's method provides an optimal solution to divide pixels into two classes by calculating a threshold that maximises the variance between two groups of pixels (the *between-class* variance) [68]. Otsu's method is entirely based on calculations performed on an image's histogram, which is an easily obtainable 1-D array. For an image with  $M \times N$  pixels with  $L$  distinct intensity levels, let  $n_i$  denote the number of pixels with intensity  $i$ . The total number of image pixels is  $MN = n_0 + n_1 + n_2 + \dots + n_{L-1}$ , and the normalised image histogram has components  $p_i = \frac{n_i}{MN}$ . It follows then that:

$$\sum_{i=0}^{L-1} p_i = 1, p_i \geq 0 \quad (2.5)$$

For a threshold  $T(k), 0 < k < L - 1$ , the input image can be divided into two classes,  $C_1$  and  $C_2$ , encompassing all pixels with intensity values in the ranges  $[0, k]$ , and  $[k + 1, L - 1]$ , respectively. With this threshold  $k$ , the probability  $P_1(k)$  that a pixel assigned to class  $C_1$  (probability of  $C_1$  occurring) is given by the cumulative sum:

$$P_1(k) = \sum_{i=0}^k p_i \quad (2.6)$$

Similarly, the probability of  $C_2$  occurring is:

$$P_2(k) = \sum_{i=k+1}^{L-1} p_i = 1 - P_1(k) \quad (2.7)$$

It follows from Eq. 2.6 that the mean intensity value  $m_1(k)$  of the pixels assigned to class  $C_1$ :

$$m_1(k) = \sum_{i=0}^k iP(i/C_1) \quad (2.8)$$

where the term  $iP(i/C_1)$  is the probability of value  $i$ , given that  $i$  comes from class  $C_1$ . From Bayes' formula:

$$P(A/B) = \frac{P(B/A)P(A)}{P(B)} \quad (2.9)$$



applied to  $iP(i/C_1)$ , Eq. 2.8 becomes:

$$m_1(k) = \sum_{i=0}^k \frac{iP(C_1/i)P(i)}{P(C_1)} \quad (2.10)$$

The probability of  $C_1$  given  $i$ ,  $iP(C_1/i)$  in Eq. 2.10 is 1, since  $m_1(k)$  only considers values  $i$  from class  $C_1$ . Additionally,  $P_i$  is the probability of the  $i^{th}$  value, which is just the  $i^{th}$  component of the histogram,  $p_i$ . Finally,  $P(C_1)$  is the probability of class  $C_1$ , which is equal to  $P_1(k)$  (Eq. 2.6), so we can rewrite Eq. 2.10 as:

$$m_1(k) = \frac{1}{P_1(k)} \sum_{i=0}^k ip_i \quad (2.11)$$

Following the same reasoning, the mean intensity value of the class  $C_2$  pixels is:

$$m_2(k) = \frac{1}{P_2(k)} \sum_{i=0}^k ip_i \quad (2.12)$$

The cumulative mean intensity up to  $k$  is then given by:

$$m(k) = \sum_{i=0}^k ip_i \quad (2.13)$$

and the average, or global intensity of the entire image is:

$$m_G = \sum_{i=0}^{L-1} ip_i \quad (2.14)$$

To assess how well the threshold at level  $k$  separates pixels into two classes, the *between-class* variance can be calculated:

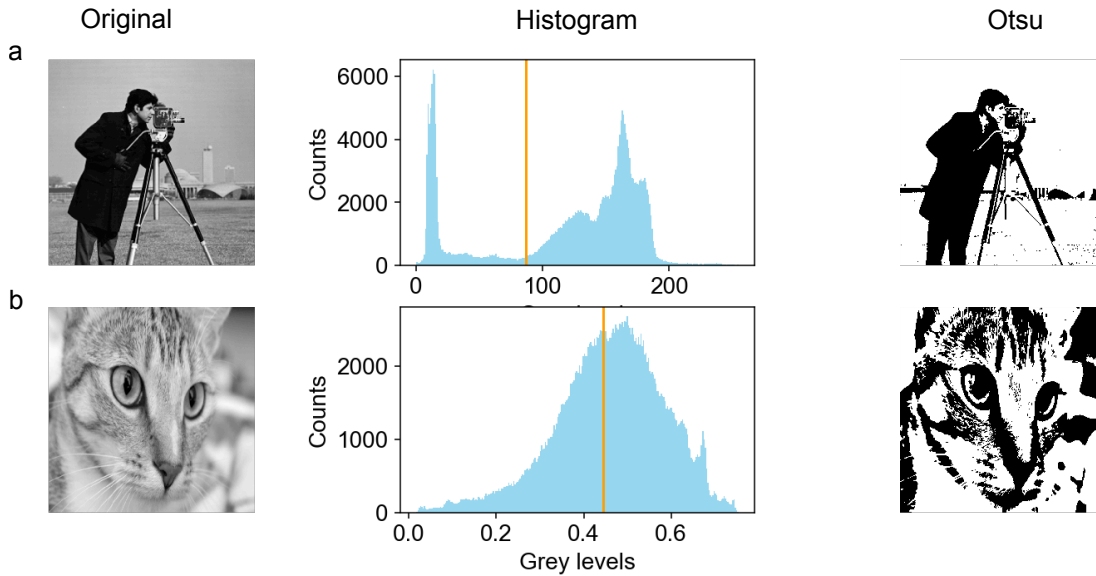
$$\sigma_B^2 = P_1(k)(m_1 - m_G)^2 + P_2(k)(m_2 - m_G)^2 \quad (2.15)$$

The more separated the two mean intensity values  $m_1$  and  $m_2$  are from each other, the larger the *between-class* variance  $\sigma_B^2$ . Therefore, algorithms implementing Otsu's method evaluate  $\sigma_B^2$  over all possible values of  $k$  to find the value  $k^*$  that maximises  $\sigma_B^2$ . If several values of  $k$  are found for which  $\sigma_B^2$  is at a maximum, then the average of the  $k$  values is used. When the optimal threshold level  $k^*$  is found, the image can be segmented as specified in Eq. 2.4:

$$f(x,y) = \begin{cases} 1, & \text{if } f(x,y) \geq k^* \\ 0, & \text{if } f(x,y) < k^* \end{cases} \quad (2.16)$$

for  $x = 0, 1, 2, \dots, M - 1$  and  $y = 0, 1, 2, \dots, N - 1$ .

Otsu's method performs best for images with a bimodal intensity distribution containing a deep valley between the two peaks, as shown in the histogram for the camera man image in Fig. 2.9 a. Images with other pixel intensity distributions, such as the cat image in Fig. 2.9 b may require local thresholding or filtering to extract features of interest such as edges, lines, and circles in the image. The derivation of Otsu's method, along with an outline for its algorithmic implementation can be found in Gonzalez and Woods, Ch.10, Section 3.3 [70]. Since Otsu's method only performs computations on 1D vectors, it is computationally very quick and can segment images in a few seconds at most.



**Fig. 2.9** Otsu's threshold (orange line) applied to an image with a bimodal (a) and unimodal (b) histogram. The open-source images "camera man", and "Chelsea the cat" from the Sci-kit image data module were thresholded using Otsu's method for this figure.

**Weka segmentation** As an alternative to traditional intensity-based image segmentation, the Trainable Weka Segmentation (TWS) plugin in ImageJ was created by Arganda-Carreras et al. [69] to combine machine learning algorithms with selected image features to generate pixel-based segmentations. In TWS segmentation is also treated as a pixel classification problem. Users can annotate images using a graphical user interface in ImageJ, adding traces of pixel regions to the specified classes. The annotated images are used as a training set for a

classifier which, once trained, can be used to categorise the remainder of the data set, or a completely new set of images. Any supervised classification or regression algorithm from the Waikato Environment for Knowledge Analysis (Weka) [71] can be used as a classifier.

Conventional segmentation methods such as Otsu's method [68] use the intensity or spatial relationship among pixels to discern regions of interest. Conversely, manual segmentation by humans leverages knowledge about the shapes, textures, local intensity gradients, etc. in images. TWS aims to leverage part of that user knowledge to generate segmentations using a rich space of image features, most of which are extracted using common filters before being used in the classifier. TWS was implemented in Chapter 4 to distinguish fluorescent puncta of HIV proteins and the associated ESCRT proteins from background signals at the cell membrane. The filters to extract features include:

- Edge detectors to indicate boundaries of objects in an image (e.g. Sobel filter).
- Noise reduction filters (e.g. Gaussian blur filter).
- Membrane detectors which enhance membrane-like features in an image using directional filters.

The default classifier in TWS, the Fast Random Forest algorithm, was used in this thesis. This classifier is based on the Random Forest, a supervised machine learning algorithm which constructs multiple decision trees based on input features, and outputs the mode of the resulting classes from all the trees. Decision trees for classification are models which repeatedly split an input data source (the root of the tree) into subsets based on input features at each node (branches), until all node results have been assigned a class label (leaves). By pooling together the most popular classification results from multiple decision trees, and using random input features for each node, the Random Forest algorithm provides a robust, accurate classification method. In TWS, the default classifier uses 200 decision trees with 2 random input features per node. The details of the Fast Random Forest algorithm are beyond the scope of this thesis; the reader can find a detailed introduction to Random Forests in Chapter 15 of Hastie et al. [72]. A detailed description of the available features and classifiers for the TWS can be found in the plugin repository [73]. For the work performed in this thesis in Chapter 4 in which less than 10 images were segmented per experimental condition, the TWS annotation for training could take between 5-10 minutes to complete, and the training itself required less than 10 seconds to produce a segmented data set.

**Morphological operations** Morphology describes pixel-wise operations, often on binary images, based on shapes of features in the image. Morphological operations (MOs) work

by probing an image with a small shape or template called a structuring element. The structuring element is scanned over all pixels in the image and compared to the corresponding neighbourhood of pixels. MOs examine if the structuring element fits within, intersects with, or neither fits nor intersects with the neighbourhood of non-zero value pixels. A new binary image is then created with non-zero values in the pixels in which the structuring element fulfilled the tested condition (fit, intersection, etc.). Structural elements in MOs are analogous to convolution kernels in digital image filtering. The two fundamental MOs are erosion and dilation. In these operations, a layer of pixels is either stripped away or added, respectively, from the inner and outer boundaries of the image. Erosion removes small elements in an image and enlarges the gaps between regions but reduces the size of the regions of interest. Dilation, on the other hand, fills in small holes and gaps between regions. Both erosion and dilation were applied in this thesis to remove unwanted signals or elements in images, or to generate accurate masks around bright regions of interest.

This concludes the brief introduction to the segmentation techniques used in the first part of this thesis. A comprehensive covering of image segmentation methods can be found in Gonzalez and Woods, Ch.10 [70].

### **2.3.2 Colocalisation analysis**

Fluorescence image colocalisation analysis is a widespread method to search for clues of interaction or association between two molecules or structures in biological samples. If two or more structures are associated or interact, they are likely to have overlapping spatial distributions. However, the reverse is not true; co-presence does not guarantee association or molecular interaction [5]. Even if two probes appear in the same location, this could either be due to random chance or to being pushed by the crowded field of unlabelled structures in the sample which are invisible to the microscope [5]. Only nearfield techniques such as FRET [74] can directly measure probe-probe interactions. Nonetheless, the repeated colocalisation of two probes in multiple samples/cell regions increases the confidence that the two could be potentially associated or interacting [75], and allow researchers to infer such interactions. Colocalisation analyses are widely used because they can be easily applied on existing multi-colour fluorescence images obtained using standard microscopes and fluorescent probes.

#### **Conventional colocalisation measures**

Colocalisation can be divided into two different measurements: correlation and co-occurrence [75]. Correlation measures the intensity relationship between two overlapping fluorophore

distributions. It can be used to compare the relative abundance between two species, or to find a functional or stoichiometric relationship between species. Co-occurrence measures the extent of spatial overlap or proximity between two fluorescent distributions, generally using object segmentation methods. In Chapters 3 and 4, the colocalisation analyses specifically measured the co-occurrence of fluorescent distributions, therefore unless specified, the terms colocalisation and co-occurrence will be used interchangeably. An overview of the methods used to test colocalisation now follows, in order of increasing complexity.

**Dye overlap measurement** The simplest implementation of colocalisation is to measure the area of signal overlap between two images. In Fig. 2.10, the images of two simulated blob patterns were binarised using Otsu's method [68] to select the shapes of the blobs from the background; the intersection and the union can then be used to calculate the degree of overlap between the two different channels as given by the following equations:

$$\frac{\text{Area}_{\text{intersect}}}{\text{Area}_{\text{Ch}_1}} = 0.24 \quad (2.17)$$

$$\frac{\text{Area}_{\text{intersect}}}{\text{Area}_{\text{Ch}_2}} = 0.19 \quad (2.18)$$

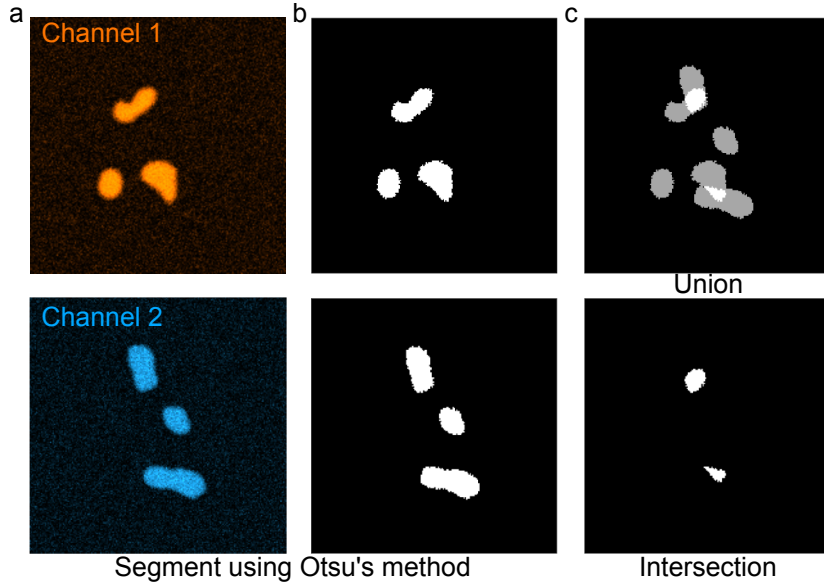
$$\frac{\text{Area}_{\text{intersect}}}{\text{Area}_{\text{union}}} = 0.12 \quad (2.19)$$

where the fractional overlap of channel 1 with channel 2 is given by the quotient of the intersection and channel 1 areas (Eqs. 2.17), and vice versa (Eq. 2.18). An overall overlap metric can be expressed as the fraction of the intersection over the union (total area) (Eq. 2.19).

This area overlap analysis can be expanded, as the segmentation step applied in Fig. 2.10 b allows discrete objects (contiguous pixel areas) to be identified with specific properties such as centroid, area, ellipticity, etc. These properties can be leveraged to compare the relative proximity of objects between colour channels. The nearest-neighbour algorithm [76] is a common implementation of this object-based analysis, which considers two objects to be colocalised if the distance between their centroids falls within a threshold distance given by:

$$d_{12} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (2.20)$$

where  $d_{12}$  is the distance between a spot with centre  $(x_1, y_1)$  in channel 1, and a spot with centre  $(x_2, y_2)$  in channel 2. A nearest-neighbour algorithm was implemented in Chapter 3 to



**Fig. 2.10** The area overlap between two colour channels is the simplest form of measuring colocalisation. (a) Two images of blobs with differing intensities and noise were simulated in ImageJ to represent two colour channels in a microscopy experiment. (b) The images were segmented using Otsu's method, and their union and intersection were computed (c).

apply a density filter on localisation data, and in Chapter 4 to compare the proximity between objects from two colour channels in different experimental conditions.

**Manders' coefficients** The dye overlap and object-based analyses described in the previous section do not take into account the signal densities of the examined fluorophore distributions. Manders et al. [77] devised a method which accounts for signal intensities when calculating the overlap between colour channels [77] in the form of two coefficients  $M_1$  and  $M_2$ :

$$M_1 = \frac{\sum_{i=1}^n x_{i_{coloc}}}{\sum_{i=1}^n x_i} \quad (2.21)$$

where

$$x_{i_{coloc}} = \begin{cases} x_i, & \text{if } y_i > 0 \\ 0, & \text{if } y_i = 0 \end{cases} \quad (2.22)$$

and

$$M_2 = \frac{\sum_{i=1}^n y_{i_{coloc}}}{\sum_{i=1}^n y_i} \quad (2.23)$$

where

$$y_{i_{coloc}} = \begin{cases} y_i, & \text{if } x_i > 0 \\ 0, & \text{if } x_i = 0 \end{cases} \quad (2.24)$$

In these equations,  $x_i$  and  $y_i$  refer to the  $i^{th}$  pixel value above an intensity threshold for colour channels 1 and 2, respectively, for a total of  $n$  pixels analysed. Thus,  $x_{i_{coloc}}$  and  $y_{i_{coloc}}$  will only be non-zero when the corresponding  $x_i$  and  $y_i$  are also above the intensity threshold. The  $M_1$  coefficient therefore measures the co-occurrence fraction of colour channel 1 on channel 2, and  $M_2$  measures the reverse [77]. Manders et al. also proposed an intensity-weighted overall overlap coefficient (Manders Overlap Coefficient, or MOC) between two channels, expressed as:

$$MOC = \frac{\sum_{i=1}^n x_i y_i}{\sqrt{\sum_{i=1}^n x_i^2} \sqrt{\sum_{i=1}^n y_i^2}} \quad (2.25)$$

where  $x_i$  and  $y_i$  are the same as defined above. The  $MOC$  is an improvement over the area overlap calculation described previously, as it takes into account intensity per area (signal density), giving larger weight to bright pixels than to dim background pixels. This causes the  $MOC$  to be susceptible to large background signal contributions, such as those originating from high non-specific labelling, strong autofluorescence, saturated pixels, or bright out-of-focus contributions [77]. These non-biologically relevant sources can bias the  $MOC$  to appear larger than its true value. Adequate thresholding, background subtraction, and/or masking are necessary to remove unwanted signal contributions and minimise their effect on the  $MOC$  [78]. Notwithstanding, the  $MOC$  is robust against variations in signal-to-noise ratio (SNR) [77], as random pixel intensity fluctuations due to noise have little effect when summing over a large number of pixels to calculate  $M_1$ ,  $M_2$ , and the  $MOC$  [78]. Manders coefficients operate on the image as a matrix of pixels, and can therefore be computed within less than a second for a simple image.

**Testing colocalisation against a random chance distribution of fluorophores** When analysing the  $MOC$ , it is important to determine whether or not the observed co-occurrence is due to a random chance distribution of the labelled species. A simple, yet coarse method to test this was proposed by Dunn [75] in which one of the colour channel images is rotated 90° to act as a negative control, and its  $MOC$  is calculated with each of the original images. A more sophisticated method was proposed by Costes [79] in which PSF-sized groups of neighbouring pixels are randomly scrambled in one of the colour channels to simulate a random distribution of fluorophores. The  $MOC$  is then calculated between the scrambled

and the unaltered channel. This process of scrambling and calculating an MOC value is repeated over ~100 to 1000 iterations to obtain a confidence interval for the mean MOC value in the case of a pseudo-random distribution of markers. If the MOC value from the experimental data exceeds 95% of the MOC values from the randomisation trials, then the colocalisation is deemed statistically significant and unlikely to be due to random chance [78]. Costes' method, although more computationally intensive, is statistically more robust than Dunn's method [78]. For a single 256x256 pixel<sup>2</sup> region (which is the size of all images acquired in Part I of this thesis), Costes' method can take up to a minute to compute for 1000 iterations, whereas Dunn's method can be computed within a second for a single image. A more detailed review on different colocalisation methods and their application can be found in Bolte and Cordelieres [80], Cordelieres and Bolte [81], Dunn et al. [75], and Aaron et al. [78].

### Coordinate-based colocalisation analysis

The emergence of super-resolution microscopy changed the way researchers think about colocalisation, since structures that previously appeared to overlap in diffraction-limited images could be revealed to be in distinct, separate locations in space [82], as demonstrated with the localised bead images in Fig. 2.8. The emergence of SMLM methods in particular, which can distinguish single molecules which are ~10-20 nm apart, led scientists to reassess how to infer interactions between sub-cellular structures [5]. The proximity threshold to determine colocalisation between two species then depends on the biological question.

Traditional colocalisation methods based on intensity overlap or correlation depend on the microscope PSF and are therefore generally not applicable to localisation data [83], however they can be applied to reconstructed SMLM images in which the pixel values in an image represent the binned number of localisations within that pixel [84]. As mentioned in section 2.2.3, the output from an SMLM experiment is not a conventional image, i.e. a pixelated array of intensity values. Rather, it is a table of  $x, y$  coordinates of individual molecules, each with an associated uncertainty and signal density, and can therefore be treated as a spatial point pattern. Colocalisation analysis becomes then the search for spatial associations between point patterns originating from different fluorophore distributions. Spatial statistics and correlation analysis methods have been developed to extract information about the spatial association (coupling or tethering) of localisations based on their relative positions.

**Spatial statistics** Spatial descriptive statistics are often used to summarise a point pattern, test hypotheses about the pattern, and fit models [85]. In spatial statistics, moments are used to describe quantitative parameters about a probability distribution. The first moment



property of a spatial point pattern is the number of points per area, and the second moment property is the expected number of points  $N$  within a distance  $r$  of another point. Ripley's K-function [86], a common spatial statistic tool for analysing point patterns, is the second moment property normalised by the number of points per area  $\lambda$ :

$$K(r) = \frac{1}{n} \sum_{i=1}^n \frac{N_{p_i}(r)}{\lambda} \quad (2.26)$$

where  $p_i$  is the  $i^{th}$  point and the sum is taken over  $n$  points. The expected  $K(r)$  value for a random Poisson distribution is  $\pi r^2$ , also defined as complete spatial randomness (CSR) [85]. Any deviations from this expectation indicate either clustering or dispersion for a point pattern. Besag [87] proposed a normalisation of the K-function so that its expected value for CSR is linear with  $r$ :

$$L(r) = \sqrt{\frac{K(r)}{\pi}} \quad (2.27)$$

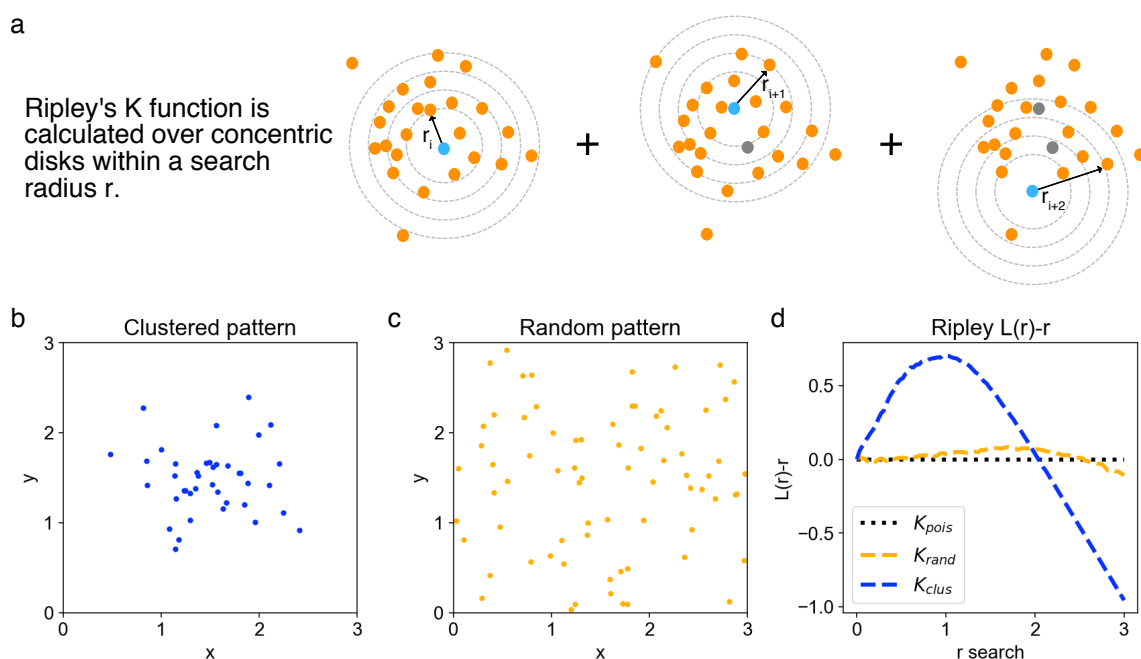
Besag's function was further normalised so that the expected value for CSR yields 0, in the form of the H-function [88]:

$$H(r) = L(r) - r \quad (2.28)$$

Ripley's K-function is most commonly used to test whether an observed point pattern is consistent with a random Poisson distribution (the null hypothesis). Positive values of the  $H(r)$  function indicate clustering over that spatial scale whereas negative values indicate dispersion, as demonstrated with a simulated clustered and random pattern in Fig. 2.11.

The value of the search radius  $r$  which maximises the value of  $H(r)$  indicates the radius of maximum aggregation, i.e. the radius of a disk centred on a sample point which on average contains the most points per area. This radius has been used as an estimate for cluster size in point patterns, however it has been reported to overestimate size of the clusters by up to a factor of twice the true radius [89]. Bivariate or multivariate generalisations of Ripley's K, L, and H-function can be used to describe relationships between two or more spatial point patterns; this is useful for performing colocalisation studies with localisation data. Depending on the density of points in the pattern at hand, as well as on the radius chosen for analysis, Ripley's K function calculation can take up to a few minutes for a single 256x256 pixel<sup>2</sup> region, and over an hour for analysing multiple fields of view ( $> 15$ ).

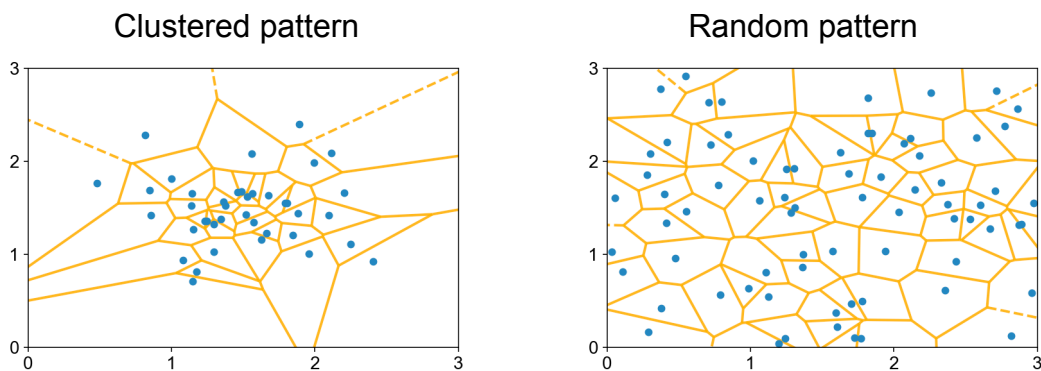
**Correlation analysis** Correlation analysis is useful to calculate spatial scales of density fluctuations in a spatial point pattern [90, 91, 92]. The Pair Correlation Function (PCF)



**Fig. 2.11** Using Ripley's K-function to analyse spatial point patterns. (a) Illustration of how Ripley's K-function is computed over concentric circles centred on each point in the pattern, for a maximum search radius  $r_n$ . (b) Uniformly clustered point pattern. (c) Random point pattern (d) Ripley's L(r)-r function computed for the clustered pattern in (b) (blue line), and the random pattern in (c) (orange line). The black dotted line represents the null hypothesis, the L(r)-r curve for a Poisson distribution. The code for this simulation can be found in the supplementary code repository for this thesis <https://github.com/pedropabloVR/supplementary-code-thesis>.

is similar to Ripley's K-function, however it measures the probability of finding another point at a distance  $R$  away from a given point, compared to that expected from a random distribution of points [91] for concentric rings instead of disks. Both Ripley's K-function and the PCF are more descriptive than a regular nearest-neighbour calculation, as they can point out properties of a spatial point pattern at different size scales e.g. regularity at a small scale and clustering at large scales. Several implementations of Ripley's K-function [89, 83, 93, 94] and correlation analysis [83, 93, 95, 96, 90, 91, 97] have been applied to measure colocalisation in SMLM data sets.

**Segmentation, followed by spatial association** Other methods apply segmentation on SMLM data sets to identify clusters, followed by spatial association metrics to measure inter-cluster proximity. One technique uses Voronoi diagrams to compartmentalise localisations into polygons and create density maps [98, 99, 100]. A Voronoi diagram is a tessellation in which a polygon corresponding to a data point represents the locus of all points in space closest to this data point [101] (Fig. 2.12). Tessellation does not require *a priori* knowledge of the clustering in the data, i.e. a search radius, and the properties of the Voronoi tiles can be readily used to describe the physical properties of clusters. A similar technique was extended to 3D data, in which convex hulls are fitted to localisation data to determine the 3D overlap of clusters [102].



**Fig. 2.12** Voronoi diagrams for the clustered and random patterns in Fig. 2.11.

A number of tools previously developed by other groups to apply some of these point pattern analysis methods are summarised in Table 2.2. In-depth reviews and analysis of the currently available colocalisation techniques for SMLM can be found in Deschout et al. [84] and Lagache et al. [83].

**Table 2.2** Software tools developed to analyse SMLM data as spatial point patterns for colocalisation and cluster analysis.

Tool	Method	Implemented in	Reference
ClusDoC	Correlation-based analysis (CBC, Malkusch et al. 2012)	Matlab	Pagoon et al. 2016 [103]
Icy-SODA	Ripley's K Function, PCF (Lagache et al. 2015)	Icy	Lagache et al. 2018 [83]
LAMA	Ripley's K Function, PCF, DBScan, OPTICS, Nearest-neighbour	Python	Malkusch et al. 2016 [96]
MOSAIC IA	Spatial interaction analysis	ImageJ/Fiji plugin	Shivandanan et al. 2013 [97]
ClusterViSU	Voronoi Tessellation	Matlab	Andronov et al. 2016 [98]
SR-Tesseler	Voronoi Tessellation	C++	Levet et al. 2015 [99]
STORM-RLA	Convex-hull fitting for 3D	Matlab	Veeraraghavan and Gourdie, 2016 [102]

## 2.4 Summary

This chapter introduced the background to the imaging and analysis methods used in Part I of this thesis. First, the historical background to the detection of single fluorescent molecules was introduced. A description of the fundamentals of localisation microscopy techniques ensued, with details on image acquisition, photoswitching, and reconstructions for *d*STORM. Subsequently, the working principles for SPT methods to analyse dynamic interactions between single molecules were presented. Finally, an introduction to chromatic offset correction, image segmentation, and colocalisation analysis methods was provided.

The following two chapters describe the individual application of these methods to examine the colocalisation of membrane-bound fluorescently labelled protein clusters at the nanoscale, in *in-vitro* models for Parkinson's Disease (Chapter 3) and HIV (Chapter 4). In each chapter, the techniques introduced here serve as building blocks for integrated image acquisition and analysis pipelines tailored to the constraints and requirements of each biological question. A number of the introduced image processing and analysis methods were implemented in newly-developed software packages; all the source code for these algorithms can be found in Github repositories and will be specified accordingly in each

chapter. Other methods were applied using existing open-source software, which will also be clearly acknowledged.

## References

- [1] Susan Jones and Janet M. Thornton. Principles of protein-protein interactions. *Proceedings of the National Academy of Sciences of the United States of America*, 93(1):13–20, 1996.
- [2] Jan A. Miernyk and Jay J. Thelen. Biochemical approaches for discovering protein-protein interactions. *Plant Journal*, 53(4):597–609, 2008.
- [3] Eric M. Phizicky and Stanley Fields. Protein-protein interactions: Methods for Detection and Analysis. *Microbiological Reviews*, 59(1):1–165, 1995.
- [4] Monya Baker. Mass spectrometry for biologists. *Nature Methods*, 7(2):157–161, 2010.
- [5] Jo A Helmuth, Grégory Paul, and Ivo F Sbalzarini. Beyond co-localization : inferring spatial interactions between sub-cellular structures from microscopy images. *BMC Bioinformatics*, 11:372, 2010.
- [6] Kota Miura and Natasa Sladoje. *Bioimage Data Analysis Workflows*. Springer, first edition, 2020.
- [7] W. E. Moerner and L. Kador. Optical detection and spectroscopy of single molecules in a solid. *Physical Review Letters*, 62(21):2535–2538, 1989.
- [8] M. Orrit and J. Bernard. Single pentacene molecules detected by fluorescence excitation in a p-terphenyl crystal. *Physical Review Letters*, 65(21):2716–2719, 1990.
- [9] E. Brooks Shera, Newton K. Seitzinger, Lloyd M. Davis, Richard A. Keller, and Steven A. Soper. Detection of single fluorescent molecules. *Chemical Physics Letters*, 174(6):553–557, 1990.
- [10] Eric Betzig and Robert J Chichester. Single Molecules Observed by Near-Field Scanning Optical Microscopy. *Science*, 262(5138):1422–1425, 1993.
- [11] R M Dickson, a B Cubitt, R Y Tsien, and W E Moerner. On/off blinking and switching behaviour of single molecules of green fluorescent protein. *Nature*, 388(6640):355–358, 1997.

- [12] B. J. Schnapp, J. Gelles, and M. P. Sheetz. Nanometer-scale measurements using video light microscopy. *Cell Motility and the Cytoskeleton*, 10(1-2):47–53, 1988.
- [13] Russell E. Thompson, Daniel R. Larson, and Watt W. Webb. Precise nanometer localization analysis for individual fluorescent probes. *Biophysical Journal*, 82(5):2775–2783, 2002.
- [14] Eric Betzig. Proposed method for molecular optical imaging. *Optics letters*, 20(3):237–9, 1995.
- [15] A. Yildiz. Myosin V Walks Hand-Over-Hand: Single Fluorophore Imaging with 1.5-nm Localization. *Science*, 300(5628):2061–2065, 2003.
- [16] Maxime Dahan, Sabine Lévi, Camilla Luccardini, Philippe Rostaing, Béatrice Riveau, and Antoine Triller. Diffusion Dynamics of Glycine Receptors Revealed by Single-Quantum Dot Tracking. *Science*, 302(5644):442–445, 2003.
- [17] Matthew P Gordon, Taekjip Ha, and Paul R Selvin. Single-molecule high-resolution imaging with photobleaching. *Proceedings of the National Academy of Sciences of the United States of America*, 101(17):6462–6465, apr 2004.
- [18] Xiaohui Qu, David Wu, Laurens Mets, and Norbert F Scherer. Nanometer-localized multiple single-molecule fluorescence microscopy. *Proceedings of the National Academy of Sciences of the United States of America*, 101(31):11298–11303, aug 2004.
- [19] Keith A. Lidke, Bernd Rieger, Thomas M. Jovin, and Rainer Heintzmann. Super-resolution by localization of quantum dots using blinking statistics. *Optics Express*, 13(18):7052, 2005.
- [20] Jennifer Lippincott-Schwartz and George H Patterson. Photoactivatable fluorescent proteins for diffraction-limited and super-resolution imaging. *Trends in Cell Biology*, 19(11):555–565, nov 2009.
- [21] Taekjip Ha and Philip Tinnefeld. Photophysics of Fluorescence Probes for Single Molecule Biophysics and Super-Resolution Imaging. *Annu Rev Phys Chem*, 63(2):595–617, 2012.
- [22] S Weiss. Measuring conformational dynamics of biomolecules by single molecule fluorescence spectroscopy. *Nature structural biology*, 7(9):724–729, 2000.

- [23] Mike Heilemann, Emmanuel Margeat, Robert Kasper, Markus Sauer, and Philip Tinnefeld. Carbocyanine dyes as efficient reversible single-molecule optical switch. *Journal of the American Chemical Society*, 127(11):3801–3806, 2005.
- [24] Mark Bates, Timothy R. Blosser, and Xiaowei Zhuang. Short-range spectroscopic ruler based on a single-molecule optical switch. *Physical Review Letters*, 94(10):1–4, 2005.
- [25] Mike Heilemann, Sebastian Van De Linde, Mark Schüttpelz, Robert Kasper, Britta Seefeldt, Anindita Mukherjee, Philip Tinnefeld, and Markus Sauer. Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes. *Angewandte Chemie - International Edition*, 47(33):6172–6176, 2008.
- [26] Eric Betzig, George H Patterson, Rachid Sougrat, O Wolf Lindwasser, Scott Olenych, Juan S Bonifacino, Michael W Davidson, Jennifer Lippincott-schwartz, and Harald F Hess. Supporting Online Material for. 2006.
- [27] Samuel T. Hess, Thanu P.K. Girirajan, and Michael D. Mason. Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophysical Journal*, 91(11):4258–4272, 2006.
- [28] M.J. Rust, Mark Bates, and Xiaowei Zhuang. Stochastic optical reconstruction microscopy (STORM) provides sub-diffraction-limit image resolution. *Nat. Methods*, 3(10):793–795, 2006.
- [29] Mike Heilemann, Sebastian Van De Linde, Anindita Mukherjee, and Markus Sauer. Super-resolution imaging with small organic fluorophores. *Angewandte Chemie - International Edition*, 48(37):6903–6908, 2009.
- [30] Sebastian van de Linde, Anna Löschberger, Teresa Klein, Meike Heidbreder, Steve Wolter, Mike Heilemann, and Markus Sauer. Direct stochastic optical reconstruction microscopy with standard fluorescent probes. *Nature protocols*, 6(7):991–1009, 2011.
- [31] Jonas Fölling, Mariano Bossi, Hannes Bock, Rebecca Medda, Christian A Wurm, Birka Hein, Stefan Jakobs, Christian Eggeling, and Stefan W Hell. Fluorescence nanoscopy by ground-state depletion and single-molecule return. *Nature Methods*, 5(11):943–945, 2008.
- [32] Gregory Giannone, Eric Hosy, Florian Levet, Audrey Constals, Katrin Schulze, Alexander I. Sobolevsky, Michael P. Rosconi, Eric Gouaux, Robert Tampe, Daniel Choquet,

- and Laurent Cognet. Dynamic superresolution imaging of endogenous proteins on living cells at ultra-high density. *Biophysical Journal*, 99(4):1303–1310, 2010.
- [33] Ralf Jungmann, Maier S. Avendano, Johannes B. Woehrstein, Mingjie Dai, William M. Shih, and Peng Yin. Multiplexed 3D cellular super-resolution imaging with DNA-PAINT and Exchange-PAINT. *Nature Methods*, 11(3):313–318, 2014.
- [34] Ralf Jungmann, Christian Steinhauer, Max Scheible, Anton Kuzyk, Philip Tinnefeld, and Friedrich C. Simmel. Single-molecule kinetics and super-resolution microscopy by fluorescence imaging of transient binding on DNA origami. *Nano Letters*, 10(11):4756–4761, 2010.
- [35] M. D. Lew, S. F. Lee, J. L. Ptacin, M. K. Lee, R. J. Twieg, L. Shapiro, and W. E. Moerner. Three-dimensional superresolution colocalization of intracellular protein superstructures and the cell surface in live *Caulobacter crescentus*. *Proceedings of the National Academy of Sciences*, 108(46):E1102–E1110, 2011.
- [36] Joerg Schnitzbauer, Maximilian T. Strauss, Thomas Schlichthaerle, Florian Schueder, and Ralf Jungmann. Super-resolution microscopy with DNA-PAINT. *Nature Protocols*, 12(6):1198–1228, 2017.
- [37] A. Sharonov and R. M. Hochstrasser. Wide-field subdiffraction imaging by accumulated binding of diffusing probes. *Proceedings of the National Academy of Sciences*, 103(50):18911–18916, 2006.
- [38] Stefan Hell. Far-Field Optical Nanoscopy. *Cell*, 2(May):1153–1158, 2007.
- [39] Stefan W. Hell and Jan Wichmann. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Optics Letters*, 19(11):780, 1994.
- [40] Markus Sauer and Mike Heilemann. Single-Molecule Localization Microscopy in Eukaryotes. *Chemical Reviews*, 117(11):7478–7509, 2017.
- [41] Hendrik Deschout, Francesca Cella Zanacchi, Michael Mlodzianoski, Alberto Diaspro, Joerg Bewersdorf, Samuel T Hess, and Kevin Braeckmans. Precisely and accurately localizing single emitters in fluorescence microscopy. *Nature methods*, 11(3):253–66, 2014.
- [42] Honglin Li and Joshua C. Vaughan. Switchable Fluorophores for Single-Molecule Localization Microscopy. *Chemical Reviews*, 2018.



- [43] Pavel Krizek, Ivan Raska, and Guy M Hagen. Minimizing detection errors in single molecule localization microscopy. *Optics express*, 19(4):155–157, 2011.
- [44] Kim I. Mortensen, L. Stirling Churchman, James A. Spudich, and Henrik Flyvbjerg. Optimized localization analysis for single-molecule tracking and super-resolution microscopy. *Nature Methods*, 7(5):377–381, 2010.
- [45] Steve Wolter, Anna Löschberger, Thorge Holm, Sarah Aufmkolk, Marie-Christine Dabauvalle, Sebastian van de Linde, and Markus Sauer. rapidSTORM: accurate, fast open-source software for localization microscopy. *Nature Methods*, 9:1040, nov 2012.
- [46] Martin Ovesný, Pavel Křížek, Josef Borkovec, Zdeněk Švindrych, and Guy M. Hagen. ThunderSTORM: A comprehensive ImageJ plug-in for PALM and STORM data analysis and super-resolution imaging. *Bioinformatics*, 30(16):2389–2390, 2014.
- [47] Daniel Sage, Hagai Kirshner, Thomas Pengo, Nico Stuurman, Junhong Min, Suliana Manley, and Michael Unser. Quantitative evaluation of software packages for single-molecule localization microscopy. *Nature Methods*, 2015.
- [48] Travis J. Gould, Samuel T. Hess, and Joerg Bewersdorf. Optical Nanoscopy: From Acquisition to Analysis. *Annual Review of Biomedical Engineering*, 14(1):231–254, 2012.
- [49] C.E. Shannon. A Mathematical Theory of Communication. *The Bell System Technical Journal*, 27(3):379–423, 1948.
- [50] Raimund J. Ober, Sripad Ram, and E. Sally Ward. Localization Accuracy in Single-Molecule Microscopy. *Biophysical Journal*, 86(2):1185–1200, 2004.
- [51] Jerry Chao, E. Sally Ward, and Raimund J Ober. Fisher information theory for parameter estimation in single molecule microscopy: tutorial. *Journal of the Optical Society of America A*, 33(7):B36, 2016.
- [52] Niccolò Banterle, Khanh Huy Bui, Edward A. Lemke, and Martin Beck. Fourier ring correlation as a resolution criterion for super-resolution microscopy. *Journal of Structural Biology*, 183(3):363–367, 2013.
- [53] Robert P.J. Nieuwenhuizen, Keith A. Lidke, Mark Bates, Daniela Leyton Puig, David Grünwald, Sjoerd Stallinga, and Bernd Rieger. Measuring image resolution in optical nanoscopy. *Nature Methods*, 10(6):557–562, 2013.

- [54] Jan Vogelsang, Thorben Cordes, Carsten Forthmann, Christian Steinhauer, and Philip Tinnefeld. Controlling the fluorescence of ordinary oxazine dyes for single-molecule switching and superresolution microscopy. *Proceedings of the National Academy of Sciences of the United States of America*, 106(20):8107–8112, 2009.
- [55] Graham T. Dempsey, Joshua C. Vaughan, Kok Hao Chen, Mark Bates, and Xiaowei Zhuang. Evaluation of fluorophores for optimal performance in localization-based super-resolution imaging. *Nature Methods*, 8(12):1027–1040, 2011.
- [56] Mike Heilemann, Peter Dedecker, Johan Hofkens, and Markus Sauer. Photoswitches: Key molecules for subdiffraction-resolution fluorescence imaging and molecular quantification. *Laser and Photonics Reviews*, 3(1-2):180–202, 2009.
- [57] Boerries Brandenburg and Xiaowei Zhuang. Virus trafficking - Learning from single-virus tracking. *Nature Reviews Microbiology*, 5(3):197–208, 2007.
- [58] Carlo Manzo and Maria F. Garcia-Parajo. A review of progress in single particle tracking: From methods to biophysical insights. *Reports on Progress in Physics*, 78(12), 2015.
- [59] Jean Yves Tinevez, Nick Perry, Johannes Schindelin, Genevieve M. Hoopes, Gregory D. Reynolds, Emmanuel Laplantine, Sebastian Y. Bednarek, Spencer L. Shorte, and Kevin W. Eliceiri. TrackMate: An open and extensible platform for single-particle tracking. *Methods*, 115:80–90, 2017.
- [60] Ingemar J. Cox and Sunita L. Hingorani. An efficient implementation of Reid’s multiple hypothesis tracking algorithm and its evaluation for the purpose of visual tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(2):148–150, 1996.
- [61] Khuloud Jaqaman, Dinah Loerke, Marcel Mettlen, Hirotaka Kuwata, Sergio Grinstein, Sandra L. Schmid, and Gaudenz Danuser. Robust single-particle tracking in live-cell time-lapse sequences. *Nature Methods*, 5(8):695–702, 2008.
- [62] I. F. Sbalzarini and P. Koumoutsakos. Feature point tracking and trajectory analysis for video imaging in cell biology. *Journal of Structural Biology*, 151(2):182–195, 2005.
- [63] Hao Shen, Lawrence J. Tauzin, Rashad Baiyasi, Wenxiao Wang, Nicholas Moringo, Bo Shuang, and Christy F. Landes. Single Particle Tracking: From Theory to Biophysical Applications. *Chemical Reviews*, 117(11):7331–7376, 2017.

- [64] Warren J Smith. Chapter 3: Aberrations. In *Modern Optical Engineering*, chapter 3, pages 61–89. McGraw-Hill, 3rd edition, 2000.
- [65] Alexander Egner and Stefan W Hell. Aberrations in Confocal and Multi-Photon Fluorescence Microscopy Induced by Refractive Index Mismatch. In James Pawley, editor, *Handbook of Biological Confocal Microscopy*, chapter 20, pages 404–413. Springer New York, 3rd edition, 2006.
- [66] Paolo Annibale, Marco Scarselli, Mattia Greco, and Aleksandra Radenovic. Identification of the factors affecting co-localization precision for quantitative multicolor localization microscopy. *Optical Nanoscopy*, 1(1):1–13, 2012.
- [67] Miklos Erdelyi, Eric Rees, Daniel Metcalf, Gabriele S. Kaminski Schierle, Laszlo Dudas, Jozsef Sinko, Alex E. Knight, and Clemens F. Kaminski. Correcting chromatic offset in multicolor super-resolution localization microscopy. *Optics Express*, 21(9):10978, 2013.
- [68] Nobuyuki Otsu. A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, C(1):62–66, 1979.
- [69] Ignacio Arganda-Carreras, Verena Kaynig, Curtis Rueden, Kevin W Eliceiri, Johannes Schindelin, Albert Cardona, and H Sebastian Seung. Trainable Weka Segmentation : a machine learning tool for microscopy pixel classification. *Bioinformatics*, 33(March):2424–2426, 2017.
- [70] Rafael C. Gonzalez and Richard E. Woods. Image Segmentation. In *Digital Image Processing*, chapter 10, pages 689–793. Prentice Hall, Upper Saddle River, 3rd edition, 2008.
- [71] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H. Witten. The weka data mining software: An update. *SIGKDD Explor. Newsl.*, 11(1):10–18, November 2009.
- [72] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. Random Forests. In *The Elements of Statistical Learning*, chapter 15, pages 587–603. Springer, 2nd edition, 2009.
- [73] Ignacio Arganda-Carreras. Trainable Weka Segmentation. [https://imagej.net/Trainable\\_Weka\\_Segmentation](https://imagej.net/Trainable_Weka_Segmentation), 2017.

- [74] Kevin; Truong and Mitsuhiro Ikura. The use of FRET imaging microscopy to detect protein – protein interactions and protein conformational changes in vivo. *Current Opinion in Structural Biology*, 11(5):573–578, 2001.
- [75] Kenneth W Dunn, Malgorzata M Kamocka, and John H McDonald. A practical guide to evaluating colocalization in biological microscopy. *American Journal of Physiology-Cell Physiology*, 300(4):723–742, 2011.
- [76] E. Lachmanovich, D. E. Shvartsman, Y. Malka, C. Botvin, Y. I. Henis, and A. M. Weiss. Co-localization analysis of complex formation among membrane proteins by computerized fluorescence microscopy: Application to immunofluorescence co-patching studies. *Journal of Microscopy*, 212(2):122–131, 2003.
- [77] E.M.M; Manders, F. J. Verbeek, and J. A. Aten. Measurement of co-localization of objects in dual-colour confocal images, 1993.
- [78] Jesse S. Aaron, Aaron B. Taylor, and Teng-Leong Chew. Image co-localization – co-occurrence versus correlation. *Journal of Cell Science*, 131(3):jcs211847, 2018.
- [79] Sylvain V. Costes, Dirk Daelemans, Edward H. Cho, Zachary Dobbin, George Pavlakis, and Stephen Lockett. Automatic and quantitative measurement of protein-protein colocalization in live cells. *Biophysical Journal*, 86(6):3993–4003, 2004.
- [80] S. Bolte and F. P. Cordelières. A guided tour into subcellular colocalization analysis in light microscopy. *Journal of Microscopy*, 224(3):213–232, 2006.
- [81] Fabrice P. Cordelières and Susanne Bolte. Experimenters’ guide to colocalization studies. Finding a way through indicators and quantifiers, in practice. *Methods in Cell Biology*, 123:395–408, 2014.
- [82] Lothar Schermelleh, Peter M Carlton, Sebastian Haase, Lin Shao, Lukman Winoto, Peter Kner, Brian Burke, M Cristina Cardoso, David A Agard, Mats G L Gustafsson, Heinrich Leonhardt, and John W Sedat. Subdiffraction Multicolor Imaging of the Nuclear Periphery with 3D Structured Illumination Microscopy. *Science*, 320(June):1332–1337, 2008.
- [83] Thibault Lagache, Alexandre Grassart, Stéphane Dallongeville, Orestis Faklaris, Nathalie Sauvonnet, Alexandre Dufour, Lydia Danglot, and Jean Christophe Olivo-Marin. Mapping molecular assemblies with fluorescence microscopy and object-based spatial statistics. *Nature Communications*, 9(1):102–108, 2018.

- [84] Hendrik Deschout, Arun Shivanandan, Paolo Annibale, Marco Scarselli, and Aleksandra Radenovic. Progress in quantitative single-molecule localization microscopy. *Histochemistry and Cell Biology*, 142(1):5–17, 2014.
- [85] Philipp M. Dixon. Ripley’s K function. In Abdel H. El-Shaarawi and Walter W. Piegorsch, editors, *Encyclopedia of Environmetrics Volume 3*, pages 1796–1803. John Wiley & Sons, Ltds., Chichester, 2002.
- [86] B.D. Ripley. The Second-Order Analysis of Stationary Point Processes. *Journal of Applied Probability*, 13(2):255–266, 1976.
- [87] Julian E. Besag. Contribution to the Discussion of the Paper by Ripley (1977). *Journal of the Royal Statistical Society Series B (Methodological)*, 39(2):193–195, 1977.
- [88] Marcelo Ehrlich, Werner Boll, Antoine Van Oijen, Ramesh Hariharan, Kartik Chandran, Max L Nibert, and Tomas Kirchhausen. Endocytosis by Random Initiation and Stabilization of Clathrin-Coated Pits. 118:591–605, 2004.
- [89] Maria A. Kiskowski, John F. Hancock, and Anne K. Kenworthy. On the use of Ripley’s K-function and its derivatives to analyze domain size. *Biophysical Journal*, 97(4):1095–1103, 2009.
- [90] Joerg Schnitzbauer, Yina Wang, Shijie Zhao, Matthew Bakalar, Tulip Nuwal, Baohui Chen, and Bo Huang. Correlation analysis framework for localization-based superresolution microscopy. *Proceedings of the National Academy of Sciences of the United States of America*, 115(13):3219–3224, 2018.
- [91] Prabuddha Sengupta, Tijana Jovanovic-Talisman, Dunja Skoko, Malte Renz, Sarah L. Veatch, and Jennifer Lippincott-Schwartz. Probing protein heterogeneity in the plasma membrane using PALM and pair correlation analysis. *Nature Methods*, 8(11):969–975, 2011.
- [92] Sarah L. Veatch, Pietro Cicuta, Prabuddha Sengupta, Aurelia Honerkamp-Smith, David Holowka, and Barbara Baird. Critical fluctuations in plasma membrane vesicles. *ACS Chemical Biology*, 3(5):287–293, 2008.
- [93] Thibault Lagache, Nathalie Sauvonnnet, Lydia Danglot, and Jean Christophe Olivo-Marin. Statistical analysis of molecule colocalization in bioimaging. *Cytometry Part A*, 87(6):568–579, 2015.

- [94] Dylan M Owen, Carles Rentero, Astrid Magenau, David Williamson, Macarena Rodriguez, and Katharina Gaus. PALM imaging and cluster analysis of protein heterogeneity at the cell surface. 454(7):446–454, 2010.
- [95] Sebastian Malkusch, Ulrike Endesfelder, Justine Mondry, Márton Gelléri, Peter J. Verveer, and Mike Heilemann. Coordinate-based colocalization analysis of single-molecule localization microscopy data. *Histochemistry and Cell Biology*, 137(1):1–10, 2012.
- [96] Sebastian Malkusch and Mike Heilemann. Extracting quantitative information from single-molecule super-resolution imaging data with LAMA - LocAlization Microscopy Analyzer. *Scientific Reports*, 6(April):6–9, 2016.
- [97] Arun Shivanandan, Jayakrishnan Unnikrishnan, and Aleksandra Radenovic. On characterizing protein spatial clusters with correlation approaches. *Scientific Reports*, 6(August):1–12, 2016.
- [98] Leonid Andronov, Igor Orlov, Yves Lutz, Jean Luc Vonesch, and Bruno P. Klaholz. ClusterViSu, a method for clustering of protein complexes by Voronoi tessellation in super-resolution microscopy. *Scientific Reports*, 6(March):1–9, 2016.
- [99] Florian Levet, Eric Hosy, Adel Kechkar, Corey Butler, Anne Beghin, Daniel Choquet, and Jean Baptiste Sibarita. SR-Tesseler: A method to segment and quantify localization-based super-resolution microscopy data. *Nature Methods*, 12(11):1065–1071, 2015.
- [100] Florian Levet, Guillaume Julien, Rémi Galland, Corey Butler, Anne Beghin, Anaël Chazeau, Philipp Hoess, Jonas Ries, Grégory Giannone, and Jean Baptiste Sibarita. A tessellation-based colocalization analysis approach for single-molecule localization microscopy. *Nature Communications*, 10(1):1–12, 2019.
- [101] Georges Voronoi. Nouvelles applications des paramètres continus à la théorie des formes quadratiques. Premier mémoire. Sur quelques propriétés des formes quadratiques positives parfaites. *Journal für die reine und angewandte Mathematik*, 133:97–178, 1908.
- [102] Rengasayee Veeraraghavan and Robert G. Gourdie. Stochastic optical reconstruction microscopy-based relative localization analysis (STORM-RLA) for quantitative nanoscale assessment of spatial protein organization. *Molecular Biology of the Cell*, 27(22):3583–3590, 2016.

- 
- [103] Sophie V. Paeon, Philip R. Nicovich, Mahdie Mollazade, Thibault Tabarin, and Katharina Gaus. Clus-DoC: A combined cluster detection and colocalization analysis for single-molecule localization microscopy data. *Molecular Biology of the Cell*, 27(22):3627–3636, 2016.





## Chapter 3

# A single-molecule method to study the role of alpha-synuclein at the synapse

Alpha-synuclein (a-syn) is a protein whose dysfunction is associated with Parkinson's Disease and is primarily localised at the pre-synapse of neurons. The pre-synapse contains vesicles ~40 nm in diameter, which enable communication between neurons by fusing to the plasma membrane to release neurotransmitters. A-syn is proposed to have role in vesicle recycling, however it is unclear if it is involved in SNARE/clathrin-mediated (CM) exo/endocytosis or in non-canonical mechanisms such as kiss-and-run, in which a vesicle transiently opens and closes a small pore when in contact with the plasma membrane [1]. In this study we used *d*STORM to clarify a-syn's spatial relationship to endocytosed vesicles in purified synaptic boutons, called synaptosomes. The proximity of VAMP-2, a protein linked to SNARE/CM mediated exo/endocytosis, to internalised vesicles was used as a control, and compared to that measured for a-syn. A software tool for colocalisation and cluster analysis of SMLM data from synaptosomes was developed, including routines for automatic synaptosome detection, chromatic registration, and density filtering to remove false positives. Ground truth models of a synaptosome were created, from which synthetic SMLM data was generated for benchmarking the colocalisation and cluster analysis scripts.

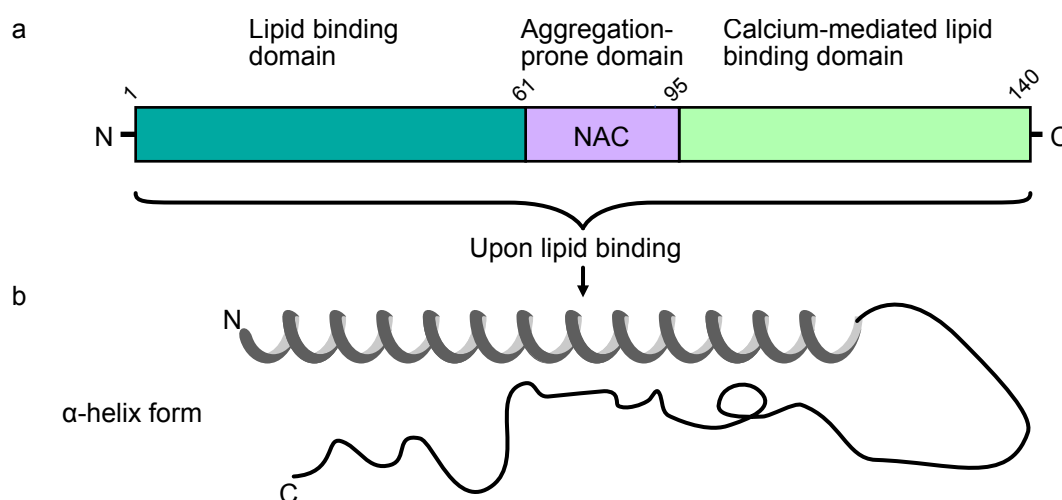
A-syn's colocalisation with internalised vesicles was found to be temperature independent, whereas VAMP-2's colocalisation with vesicles decreased significantly at sub-physiological temperatures. This suggests a-syn has a significant association with vesicles internalised via recycling mechanisms which are temperature independent and occur even when membranes are rigid, such as kiss-and-run. Furthermore, both a-syn and VAMP-2 clusters were shown to increase in size upon calcium depletion, potentially due to a link between a-syn and VAMP-2 in calcium-dependent clustering of synaptic vesicles.

## 3.1 Introduction to alpha-synuclein

In the brain, the protein a-syn is primarily localised at the pre-synapse of neurons in both a structurally disordered form and a partially ordered membrane-bound form [2]. The dysfunction of a-syn leads to the formation of protein aggregates, which in turn form the main component of so-called Lewy Bodies and Lewy Neurites in neurons. These are dense protein deposits which have been identified as the hallmarks of Parkinson's Disease [3]. A-syn is not an essential protein [2] and therefore its physiological role at the pre-synapse is unclear. The models used to study its role have led to continued scientific debate, as different cell types and expression systems have resulted in conflicting findings.

### 3.1.1 Structure and function

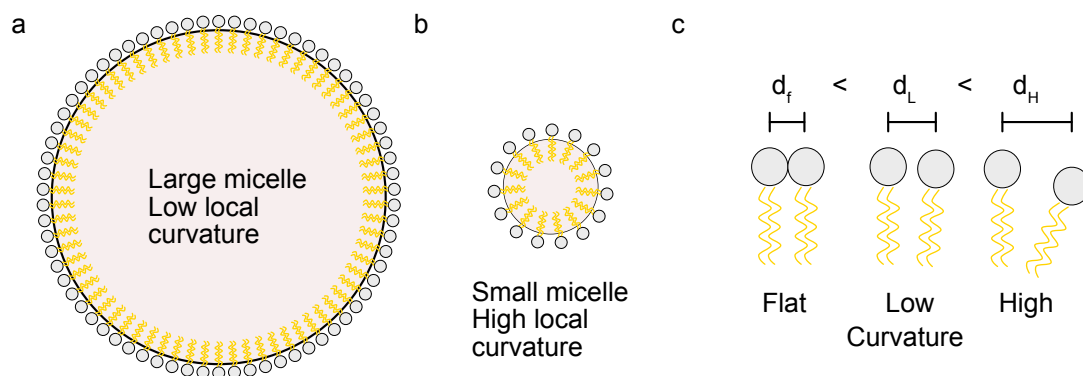
A-syn is a 140 amino acid (aa) protein composed of three distinct regions illustrated in Fig. 3.1: the N-terminus (aa 1-60) which has been shown to mediate binding to lipid membranes [4], the non-amyloid- $\beta$  component (NAC) (aa 61-95) which is hydrophobic and prone to aggregation, and the C-terminus (aa 96-140) which is highly negatively charged and can also bind lipids in a calcium-dependent manner [5].



**Fig. 3.1** Diagram of a-syn's structure showing the lipid-binding domains and the aggregation prone NAC region. (a) The N-terminus (aa 1-60) can bind lipids such as vesicles and membranes [4], and the C-terminus can also bind lipids but in a calcium-dependent manner [5]. (b) Cartoon of a-syn's helical shape after binding lipids either via its N or C terminus.

### 3.1.2 A-syn and lipid vesicles

As an unbound protein in the cytosol, a-syn is intrinsically disordered, however upon lipid binding a-syn becomes  $\alpha$ -helical [6] as illustrated by the cartoon in Fig. 3.1 b. Due to its N-terminus (aa 1-90) lipid-binding capacity, many studies have investigated the interaction between a-syn and synthetic lipid mono- and bilayers [7, 8, 9, 10, 11, 12, 13], and show that a-syn can sense curvature and preferentially binds to small vesicles [14, 11, 15, 16]. A-syn favours binding to highly curved surfaces such as those of small vesicles in which the spacing between each phospholipid unit is larger, as illustrated by the drawings in Fig. 3.2.

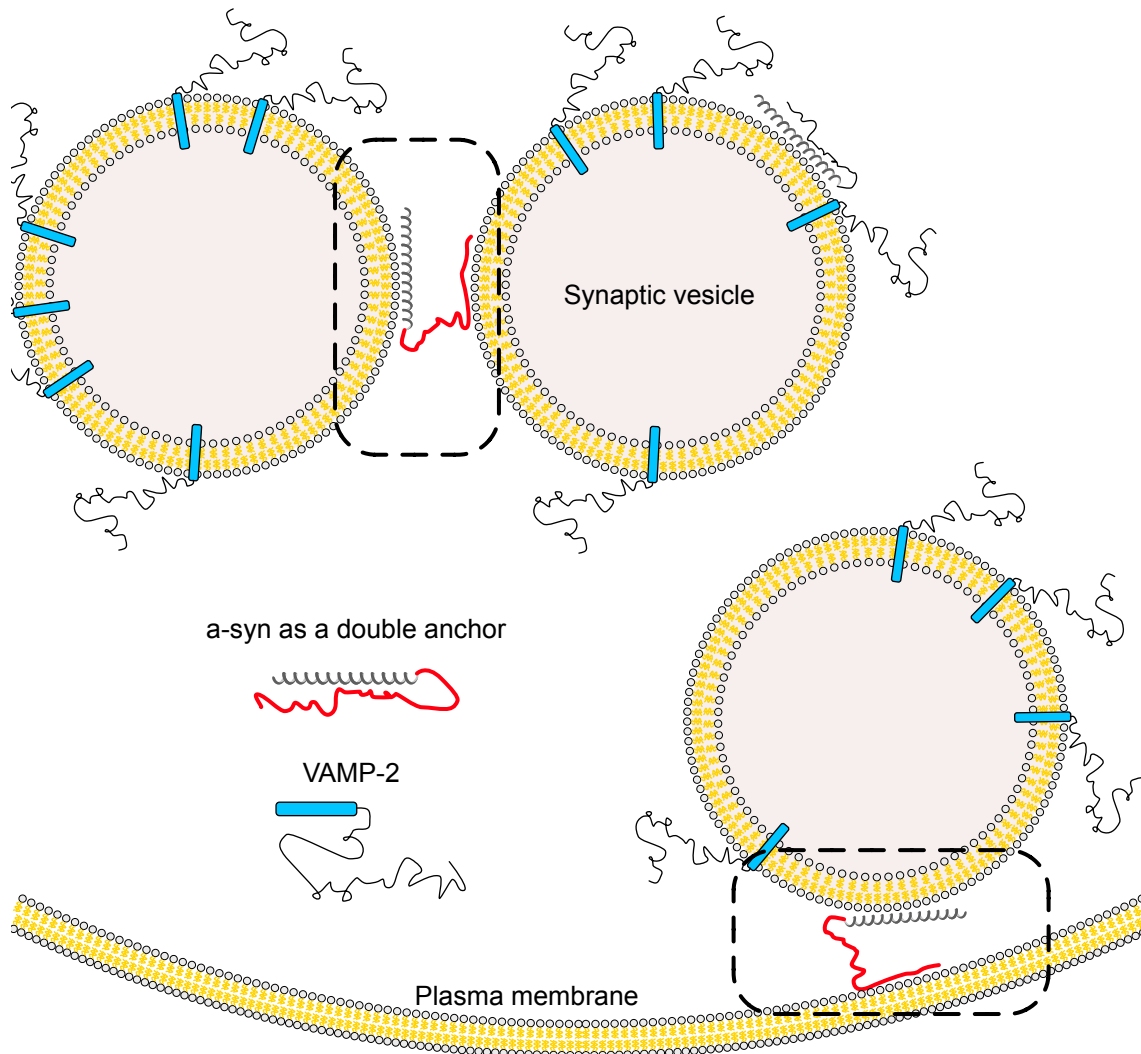


**Fig. 3.2** Cartoon of micelle-based membranes to illustrate how curvature affects packing of phospholipid groups. The spacing between adjacent phospholipid is smallest ( $d_f$  and  $d_L$ ) for surfaces with flat or low curvature (a-b), and largest ( $d_H$ ) for surfaces with high curvature (c) such as small micelles or vesicles. This allows certain proteins to “sense curvature”, i.e. to bind to membranes which have a larger spacing between their phospholipid units.

Furthermore, a series of studies from our group have recently showed that a-syn can bind physiological vesicles isolated from rat brains via its N-terminus and in a calcium-dependent manner at the C-terminus, leading to the formation of an extended double-anchor mechanism [17, 18, 5] (Fig. 3.3). This mechanism indicates a-syn can bind two vesicles together, or a vesicle to the plasma membrane (Fig. 3.3), suggesting a-syn’s physiological role could be to act as a calcium-dependent modulator of vesicle homeostasis and/or vesicle recycling [5].

### 3.1.3 A-syn and membrane trafficking

Vesicle recycling is a well-documented membrane trafficking pathway at synapses, which are the essential communication junctions between neurons. Vesicles filled with neurotransmitters dock to the plasma membrane in a region called the active zone, and then fuse with the membrane (exocytosis) releasing the neurotransmitters. Vesicle components are later



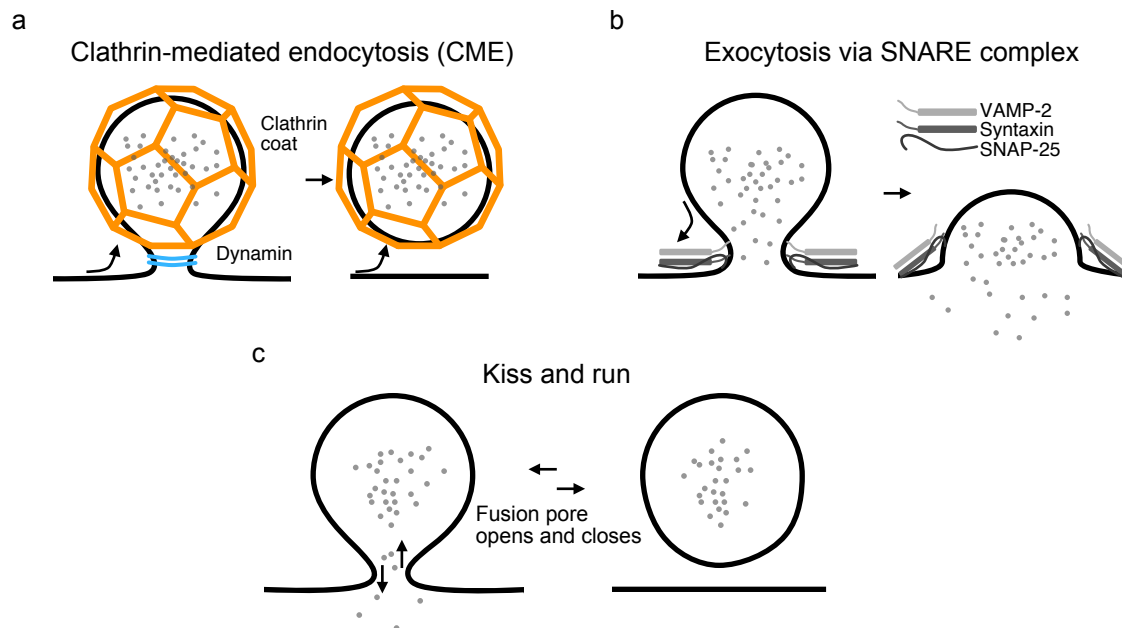
**Fig. 3.3** Drawing showing a-syn's potential double anchor mechanism between two vesicles, or a vesicle and the plasma membrane, through its lipid binding properties at its N and C termini. The dotted rectangles show how a-syn can bind either two vesicles together or a vesicle to the plasma membrane.

retrieved from the plasma membrane (endocytosis) and are refilled with neurotransmitters to repeat the cycle.

A-syn knockouts (KOs) have been previously shown to affect the size of the synaptic vesicle pool [19, 20], i.e. the vesicle group readily available for recycling, suggesting a-syn acts as a modulator of vesicle homeostasis at the presynapse. This potential modulatory behaviour, along with a-syn's ability to bind vesicles, suggests a-syn may have a role in vesicle recycling. Furthermore, a-syn is not present in all vesicles, but only in sub-populations, which suggests it might have a specific function in vesicle recycling [21].

A-syn's role, however, is difficult to pinpoint, as there are multiple mechanisms involved in vesicle recycling in neurons [22, 23]. Understanding these mechanisms is complicated by the presence of multiple pathways of exo/endocytosis, such as the canonical clathrin-mediated endocytosis (CME) (Fig. 3.4 a), SNARE regulated membrane fusion (Fig. 3.4 b), or the non-canonical activity-dependent bulk transport, fast compensatory, ultrafast, and kiss-and-run mechanisms [22] (Fig. 3.4 c). CME is a scaffolded inward budding of the plasma membrane which uses a complex coat of over 50 endocytic proteins, including clathrin, to induce membrane curvature [24] and take in neurotransmitters. SNARE regulated membrane fusion is the opposite topological process, whereby a vesicle comes into contact with the plasma membrane and a multi-protein complex known as the SNARE complex promotes the creation of a fusion pore between the two lipid membranes. Both of these canonical methods involve energy expensive processes, requiring ATP [25] to mobilise complex protein machineries, and take from 10-20 seconds to complete. In contrast, non-canonical mechanisms such as kiss-and-run, in which a fusion pore is transiently opened between a vesicle and the plasma membrane for quick uptake or release of neurotransmitters, or ultrafast endocytosis, in which a vesicle is formed independent of clathrin within 200-300 ms [22], occur in time scales of milliseconds. Specifically, kiss-and-run events have been reported to still occur at sub-physiological temperatures (e.g. 4°C), whereas canonical endocytosis mechanisms such as CME, as well as ultrafast, are inhibited at sub-physiological temperatures [25]. This temperature dependency in vesicle recycling mechanisms provides a way to observe the effect of different mechanisms by incubating samples either in physiological or sub-physiological temperatures.

Moreover, differences in cell type used, method and frequency of stimulation to induce trafficking events, and methods to observe exo/endocytosis, e.g. biochemical, fluorescence microscopy, electrophysiology, etc., make it difficult to determine the identity of membrane-trafficking mechanisms [22]. Moreover, the speed of neurotransmitter release (~millisecond scale [26]) and the size of the vesicle pores (~1-2 nm across in neuronal cell types [27]) make it difficult to image vesicle recycling events in real time.



**Fig. 3.4** Drawing showing several mechanisms of vesicle recycling. (a) Clathrin mediated endocytosis, (b) Exocytosis via the SNARE complex, and (c) kiss-and-run.

## Exocytosis

The role of a-syn in exocytosis and endocytosis is controversial, as different studies of a-syn's influence on membrane trafficking events report contrasting views. Regarding exocytosis, the modest overexpression of wild type a-syn has been shown to inhibit synaptic vesicle exocytosis in PC12 cells and chromaffin cells [28], in neurons [29, 30], and in mice [31, 32, 33]. Conversely, an increase in neurotransmitter release (exocytosis) was determined in neurons injected with a-syn [34]. A-syn knockouts (KO) on the other hand have been reported to show decreased exocytosis in neurons [34] and in mice [35, 31, 36]. Other studies using a-syn KO mice have reported both no effect on exocytosis [2, 37], as well as an increase in exocytosis [38, 39, 40].

The contrasting views presented above indicate a-syn may be implicated in an exocytosis mechanism. A-syn has been shown to bind VAMP-2, a member of the SNARE complex (Fig. 3.4 b), at its C-terminus and to promote SNARE complex assembly [41]. This link between a-syn and VAMP-2 has recently been suggested to help in the physiological clustering of synaptic vesicle pools, thereby restricting their egress to the plasma membrane and attenuating vesicle recycling and exocytosis [42]. Another recent study found overexpressed and endogenous a-syn promotes cargo discharge and reduce the closure of the fusion pore in adrenal chromaffin cells and neurons [43], suggesting a-syn has a potential role in kiss-and-run (Fig. 3.4 c).

## Endocytosis

Recent studies using models with a-syn overexpression suggest a-syn promotes clathrin-mediated endocytosis (CME) of synaptic vesicles. These studies reported an increase in membrane uptake after measuring increased internalization levels of surface receptors such as transferrin [44], NMDAR subunit NR1 [45], and dopamine transporter [46] in neurons. During CME, clathrin is the main component of the coating over the budding vesicle which supports and promotes membrane curvature, and dynamin is a key protein which coils around the neck of the budding membrane to achieve membrane scission and detachment of the vesicle [47] (Fig. 3.4 a). A study in which dynamin 1,3 were knocked out in synapses, resulting in vesicle buds that did not fully detach from the membrane, showed that a-syn and clathrin clusters colocalised with the membrane in similar abundance [20]. This study suggests a-syn potentially acts at the very beginning of CME to induce membrane curvature before fission of the vesicle by dynamin in CME [20].

The incomplete picture for the role of a-syn in synaptic vesicle recycling likely stems from the variety of cell types, stimulation techniques, and analysis methods used in the studies mentioned above which may not capture all release and recycling events. Furthermore, different expression systems of a-syn, i.e. overexpression or single, double, or triple KO can complicate interpretations. A detailed review on a-syn's potential role in vesicle recycling mechanisms is provided by Lautenschlager et al. [6].

### 3.1.4 The role of a-syn at the presynapse

#### A-syn and non-canonical vesicle trafficking mechanisms

In the context of the contrasting views presented above on the role of a-syn in synapses, we were interested in studying the role of endogenous a-syn at the presynapse to clarify its relationship to vesicle trafficking mechanisms. This is important, as a better understanding of a-syn's function at the synapse could lead to new insights on how it then misfolds and becomes part of Parkinson's Disease pathology. We examined the spatial relationship between a-syn and endocytosed vesicles and compared it to that of VAMP-2 and endocytosed vesicles. VAMP-2 is an abundant synaptic protein which is known to be involved in exocytosis as a member of the SNARE complex [41], and to be taken up during CME [23, 20]. Therefore, VAMP-2 will likely overlap with vesicles recycled via these two bulk mechanisms. We investigated if a-syn colocalised with internalised vesicles via CME or the SNARE complex like VAMP-2, or if it colocalised with vesicles internalised via different (non-canonical) mechanisms.

### **Calcium mediates vesicle recycling in dopaminergic neurons**

Dopaminergic neurons (DN) are the most vulnerable neurons in the substantia nigra pars compacta (SNc), one of the primary brain regions affected in Parkinson's Disease. Calcium is an important regulator of DN activity, as these neurons have voltage-gated calcium ( $\text{Ca}^{2+}$ ) channels which mediate autonomous pace making [48, 49] i.e. the oscillatory generation of action potentials. Calcium is also one of the key molecules involved in vesicle release and recycling along with ATP, as it provides cues for the release or uptake of neurotransmitters via synaptic vesicles [50]. Since the vesicle-binding properties of a-syn at the C terminus are mediated by calcium [5], we applied cluster analysis methods to investigate how the single-molecule distribution of a-syn was affected by the presence or absence of calcium compared to that of VAMP-2.

## **3.2 Experimental Methods**

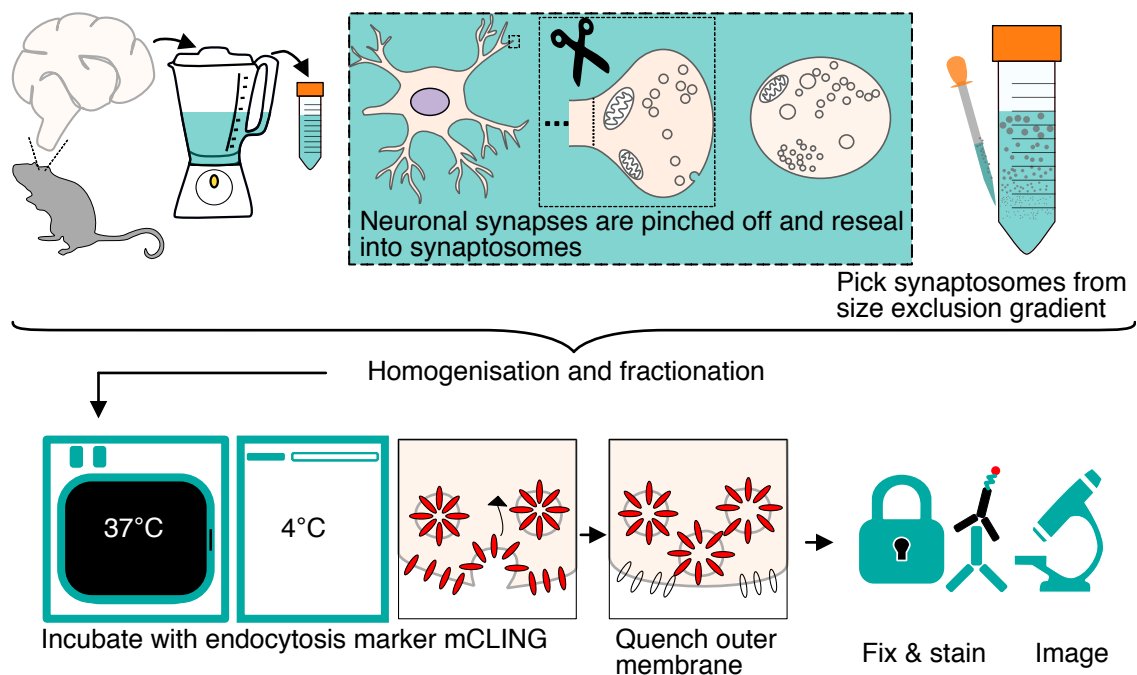
### **3.2.1 Synaptosome preparation and stimulation**

Synaptosomes were obtained by homogenising two WT Sprague-Dawley rat brains and then separating pinched-off synapses from neurons in a size-exclusion gradient. The preparation was performed by Dr. Amberley Stephens from the Molecular Neuroscience Group, and it is illustrated in Fig. 3.5. The detailed protocol for preparing synaptosomes can be found in [5].

### **3.2.2 Fluorescent labelling to visualise synaptic proteins and vesicle uptake**

Synaptosomes were stimulated for 30 min at 37°C or 4°C (Fig. 3.5) with extracellular solution containing a final concentration of 0.05 nM mCLING, which is a fixable lipid intercalating dye labelled with Atto 647N (#710 006AT1, SYNaptic SYstems, Göttingen, Germany) made at either 5 mM KCl + 2 mM  $\text{CaCl}_2$  (mimicking a physiological extra cellular buffer), 5 mM KCl + 1 mM EGTA (a calcium depleted extracellular buffer), 70 mM KCl + 1 mM EGTA (mimicking stimulation without a calcium buffer). Synaptosomes were then fixed with 4% formaldehyde and 0.2% glutaraldehyde (Sigma-Aldrich) in PBS. Fixation was quenched by washing with 0.1 M glycine in PBS for 5 min. To chemically quench the mCLING bound to the outer membrane of the plasma membrane of the synaptosomes, i.e. to eliminate its fluorescent emission and to allow us to visualise internalised mCLING, 0.75 mM bromophenol blue was added for 5 min and washed off with PBS three times [51]. Synaptosomes were stained for alpha-synuclein (D37A6 XP, 1:500 dilution, rabbit, Cell





**Fig. 3.5** Illustration of the synaptosome preparation procedure. The brain from a rat was extracted and homogenised to pinch off the synaptic ends of neurons, which reform into synaptosomes. These were incubated at 37°C and 4°C with a membrane intercalating dye (mCLING) to tag internalised vesicles, and then fixed and stained for imaging.

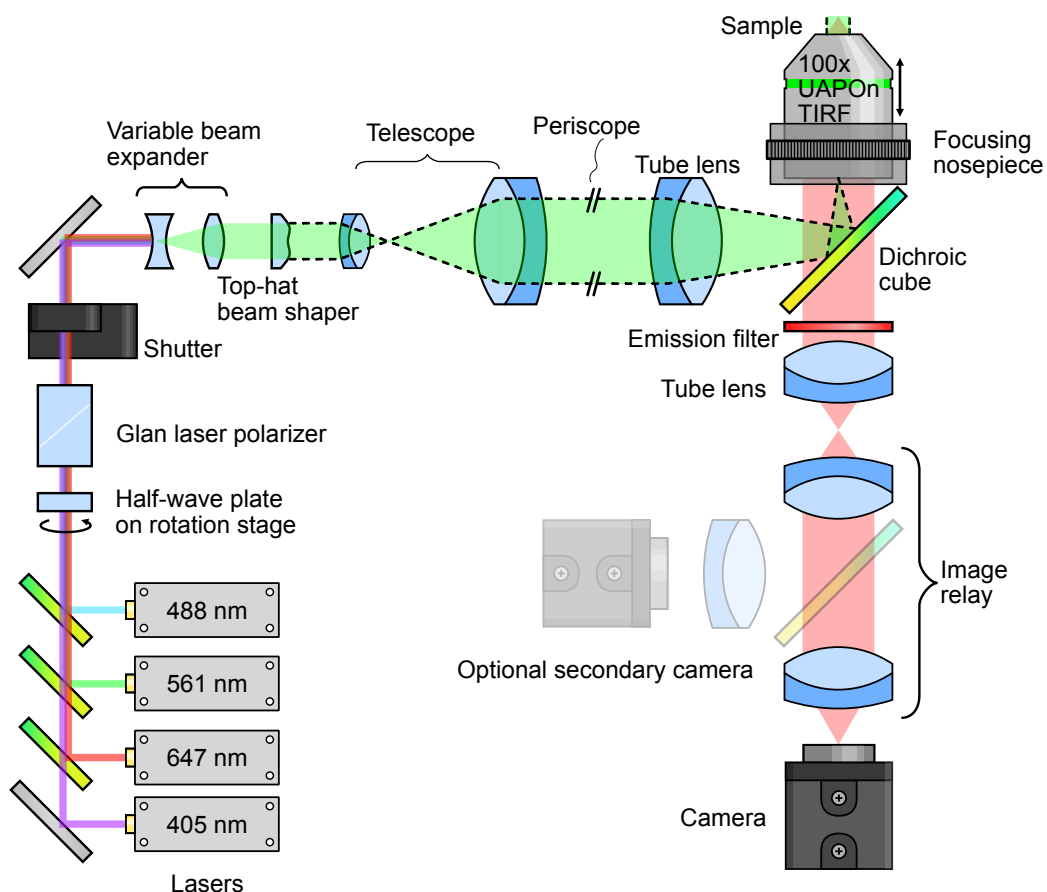
Signalling, Danvers, US) and VAMP2 (104211, 1:500 dilution, mouse, SYnaptic SYstems, Göttingen, Germany). The secondary antibodies anti-rabbit Alexa Fluor 568 (ab150067, 1:500 dilution, Abcam) and anti-mouse Alexa Fluor 488 (ab175700, 1:500 dilution, Abcam) were used to visualise protein localisation. The fluorescent staining of the synaptosomes was also performed by Dr. Amberley Stephens and is illustrated in the second row of Fig. 3.5. At this point, the synaptosome samples were plated on coverslips, ready for imaging, which is where my responsibility and contribution in the project began. Onwards, all the experiments, analysis, and simulations were performed by me unless otherwise specified.

### 3.2.3 Multi-colour *d*STORM to study synaptic proteins in synaptosomes

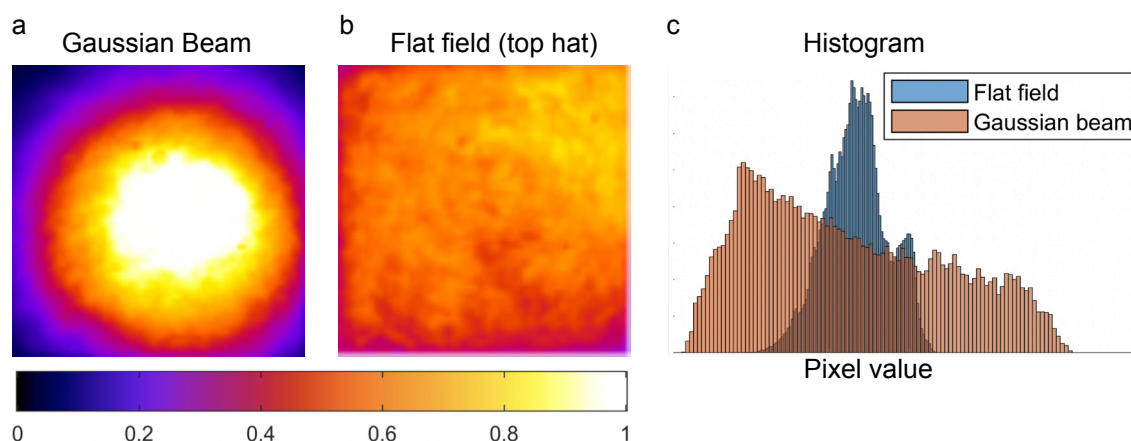
#### TIRF/HILO microscope used for *d*STORM

A custom-built wide field microscope based on an Olympus (Center Valley, PA) IX-73 frame was used to perform *d*STORM. The system included a 488-nm laser (Coherent Sapphire 488-300 CW CDRH), a 561-nm laser (Cobolt Jive 500 561 nm), and a 647-nm laser (MPB Communications Inc. VFL-P-300-647-OEM1-B1). Laser light entering the microscope frame was reflected from a dichroic mirror (Chroma ZT488/561/647rpc) into a 100x 1.49 NA oil objective lens (Olympus UAPON100XOTIRF), and then onto the sample. Light emitted by the sample passed through the dichroic and a set of 25 mm band-pass filters (Semrock FF01-525/45-25, Semrock FF01-600/37-25, and Semrock FF01-680/42-25 respectively for the 488, 561, and 647-nm laser lines) before reaching the microscope side port. Images were then relayed onto a camera (Andor iXon Ultra 897) by a 1.3x magnification Cairn Twincam image splitter. The reflective arm of the splitter was not used for these experiments. The optical layout for this setup is illustrated in the schematic shown in Fig. 3.6.

The illumination profile in most conventional microscopes, including the one showed in Fig. 3.6 is Gaussian, which implies a trade-off between illumination uniformity and photon efficiency. Typically, microscope users crop the camera field of view (FOV) to use only the center of the Gaussian beam, however this wastes precious photons at the edges of the field. Alternatively, the full FOV can be used, at the cost of under-exposing the edges. To ensure a uniform rate of fluorophore activation across the FOV during *d*STORM experiments, a Gaussian-to-top-hat beam shaper (Topag Lasertechnik GmbH GTH-5-250-4-VIS) was implemented and characterised (Fig. 3.7), as described by Rowlands et al. [52]. This top-hat beam shaper is compatible with the multi-colour implemented in this chapter, and unlike other reported beam homogenisers using micro-lens arrays [53] which can be cumbersome to align, this system only uses a single element which is easy to align. During experiments, multi-colour images were acquired sequentially, starting with the 647-nm excitation channel,



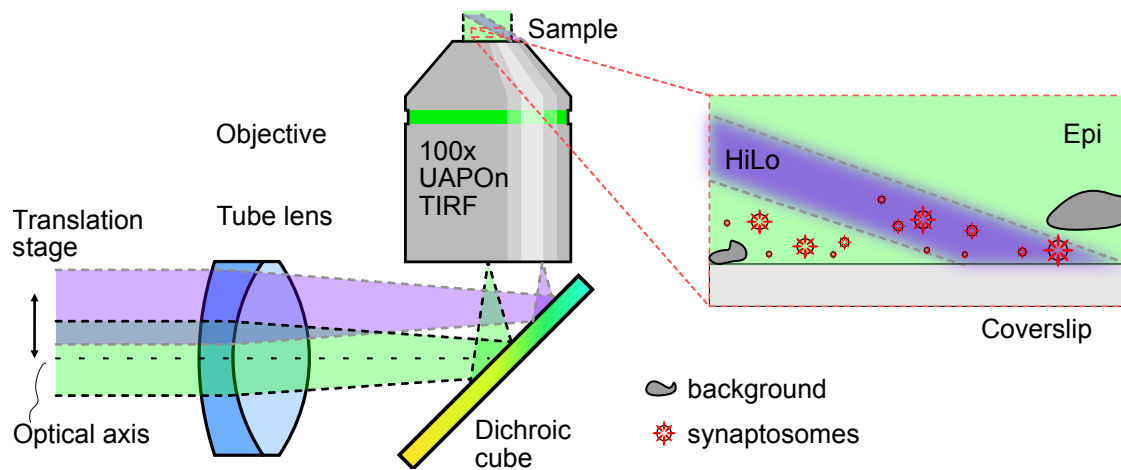
**Fig. 3.6** Optical layout of the widefield microscope used for single-molecule localisation. The layout includes 4 laser lines, a half-wave plate for intensity modulation, a gaussian-to-top hat beam shaper as described in Rowlands et al. [52], a 100X oil-immersion TIRF lens with 1.49 NA, and an electron-multiplying charge-coupled detector (EM-CCD).



**Fig. 3.7** Demonstration of the improvement in pixel uniformity when using a top hat beam shaper, compared to a Gaussian. Images in a  $60.4 \times 60.4 \mu\text{m}^2$  field of view of a Chroma test slide. The display range is identical for both the Gaussian and the top-hat image; the lower limit is the dark level, and the upper limit is  $1.2 \times$  the 95th percentile value for each image. Figure adapted from Rowlands et al. [52], a 100X oil-immersion TIRF lens with 1.49 NA, and an electron-multiplying charge-coupled detector (EM-CCD).

followed by the 561 and the 488 nm channels. This order was chosen to avoid crosstalk, namely the excitation of a dye by laser lines other than its corresponding excitation line, as well as to avoid bleaching of the far-red dye by the 488 and 568 nm laser lines. Highly-inclined illumination and laminated optical sheet (HILO) [54] was used to minimise the signal contribution from the inhomogeneous background in the synaptosomes, as this not only allows for a better signal-to-background ratio than epi-illumination, but also for a larger axial excitation range compared with total internal reflection fluorescence (TIRF) illumination (albeit with a lower signal-to-background ratio)[54]. HILO is achieved by translating the collimated beam exiting the telescope in Fig. 3.6 laterally across the tube lens, which in turn laterally shifts the excitation spot in the back focal plane of the objective towards the edge of the clear aperture. This results in a collimated, off-axis beam exiting the objective towards the sample at a steep angle, as illustrated in Fig. 3.8. The inset in the sample plane on Fig. 3.8 shows a cartoon comparing the epi illumination scheme (green) in which the entire sample is flooded with light and therefore the background is high, and the HILO scheme (purple) in which a thin sheet of light exits the objective at an angle and selectively illuminates only a slanted plane  $\sim 6\text{-}15 \mu\text{m}$  thick in the sample [54].

For the *d*STORM experiments, 16000 frames were acquired per colour channel, with an exposure time of 10 ms per frame and an electron-multiplying (EM) gain of 200. The number of frames was chosen to be the largest number of frames the camera would capture before splitting the acquired frame stacks into two, as a practical way to keep the reconstruction



**Fig. 3.8** Schematic of the HILO implementation in the microscope setup from Fig. 3.6. The epi-illumination beam is translated away from the optical axis before entering the telecentric tube lens, such that the laser focus spot is translated in the back focal plane of the objective. This results in an inclined light sheet exiting the objective towards the sample, as illustrated in the zoomed inset.

step simple. This number of frames also proved sufficient to properly sample our structures, as they were densely labelled. The exposure time was chosen as the minimal time required to capture at least 1000 photons per frame on average for the 647 nm channel, and at least 500 photons per frame on average for the 568 and 488 nm channels, determined experimentally by calculating the number of photons obtained per frame as a function of exposure time. Prior to imaging, samples were placed on the microscope stage for ~40 minutes to allow any thermal or mechanical fluctuations of the system to reach a steady-state, and therefore minimise sample drift during the acquisitions. This settling time was determined empirically as well, by visualising the drift in the image plane of the camera as a function of time and recording the time at which the focus did change over a 10-minute span. Additionally, an Olympus nosepiece (Olympus IX2-NPS) was used to decouple the objective from the mechanical stage during fine focusing to minimise mechanical and thermal drift. The irradiance at the sample plane used for the *d*STORM experiments was measured to be approximately 10 kW/cm<sup>2</sup> for all laser lines, as recommended in the seminal protocol for *d*STORM [55]. The microscope was designed and built by Dr. Romain Laine and Dr. Florian Ströhl in 2015, and it was rebuilt in 2017 by Dr. Ströhl, Dr. Christopher Rowlands, and me.

### Photoswitching buffer used for *d*STORM experiments

The adequate photo switching of fluorescent dyes for *d*STORM requires not only high laser excitation to drive the fluorophores into a dark state, but also a buffer solution with oxygen

scavengers and a reducing thiol to ensure a low duty cycle for high localisation precision, as discussed in Chapter 2. The recipe in Table 3.1 was used for all our *d*STORM experiments, adapted from Van de Linde et al. [55]:

**Table 3.1** Ingredients for photo switching buffer used in *d*STORM experiments.

Amount	Reagent	Comments
<b>Enzyme Stock Solution</b>		
100 μL	Catalase	1 mol/L
200 μL	TCEP	
25 mL	Glycerine	
22.5 mL	Distilled water	
1.25 mL	KCl	1 M
1 mL	Tris-HCl pH 7.5	1 mol/L
50 mg	Glucose oxidase	
<b>Glucose stock solution</b> (50 mL). Make 1 mL aliquots and store at -20°C		
5 g	Glucose	1 mol/L
45 mL	Distilled water	
5 mL	Glycerine	
<b>MEA stock solution</b> (10 mL). Make 200 μL aliquots and store at -20°C		
1,136 g	MEA-HCl	Dissolve for 1 mol/L soln. Adjust pH to 8.
10 mL	Distilled water	
<b>dSTORM buffer</b> (1.5 mL). Volume for single LabTek well		
543 μL	Glucose	Oxygen scavenger
69 μL	Enzymes	Oxygen scavenger
105 μL	MEA	Reducing thiol
783 μL	PBS	
24 μL	KOH	1 M for pH 8

### Reconstruction of SMLM data sets

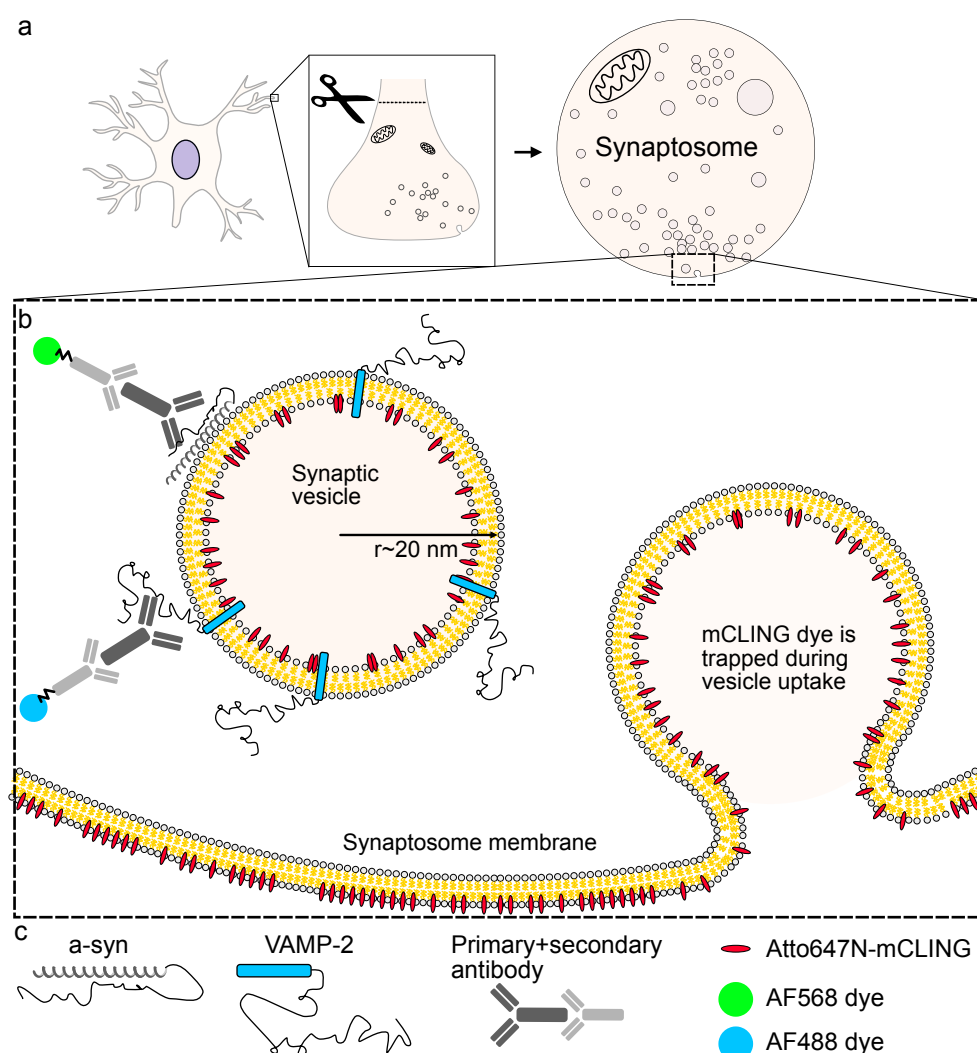
The reconstruction of the acquired frames was done in thunderSTORM [56], using the default parameters for filtering (Wavelet filter, order 3, scale 2), approximate and sub-pixel localisation of molecules (Gaussian PSF with a 3 pixel fitting radius and a 1.6 pixel initial sigma), and visualisation (Average Shifted Histograms, 10x pixel magnification). The camera parameters used were: 117 nm pixel size measured at the image plane, 15.76 photoelectrons per A/D count [57], 200 A/D count base level [57], EM Gain of 200.

### 3.2.4 Statistical analysis

The colocalisation and cluster analysis results in this chapter were presented using estimation graphics [58]. In the analysis of two data groups, individual data points are displayed in bee swarm plots, and the effect size, or difference of means between groups, is plotted on a separate axis with its origin aligned with the mean of the test group. A 95% confidence interval calculated with bootstrap resampling is plotted as a sampling distribution around the difference of means and used as an indicator of precision. This is referred to as a Gardner-Altman plot. Summary statistics (mean and standard deviation) are displayed as a gapped line to the right of each bee swarm plot. Two biological replicate experiments per temperature condition and per calcium condition were conducted, and in each replicate over 10 images were acquired and reconstructed. For each condition analysed in the results section,  $N$  represents the number of extracted and analysed synaptosomes. The number of individual localisations analysed in each data set was not quantified, since dSTORM experiments suffer from overcounting of localisations and therefore the number of localisations varies widely and is not representative of the stoichiometry of the sample, especially in dense data sets like the synaptosomes. Estimation graphics were generated using the `dabestr` package from Ho et al. [58]. The scripts for generating estimation plots from the colocalisation (`part1Ch3_results_synaptoAnalysis_colocalisation_ripleysK.R`) and cluster analysis (`part1Ch3_results_synaptoAnalysis_clusterSize_RMSD.R`) results can be found in the supplementary code repository for this thesis <https://github.com/pedropabloVR/supplementary-code-thesis>.

## 3.3 Synaptosomes as a model to study vesicle recycling

Synapses are abundant in cultured neurons, and these have been previously examined using SMLM [59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72]. However, in-vitro neuronal imaging can suffer from a strong background signal contribution from the cell body, and synapses might be partially obscured or concealed by other cellular structures. Instead, we used pinched-off synapses from rat brains, called synaptosomes (Fig. 3.9), to study the distribution of a-syn and VAMP-2 inside synapses with minimal background contribution from cellular structures. Moreover, isolating synaptosomes directly from rat brains to observe the role of a-syn is faster than culturing primary neurons from rat brains, which takes weeks. Synaptosomes are a well established model, which remain metabolically active and can be stimulated to trigger vesicle release shortly after preparation [73]. Synaptosomes can also be easily treated with drugs and differing temperature to allow observation of different endo/exocytosis mechanisms.



**Fig. 3.9** Synaptosome model to visualise colocalisation of mCLING-stained internalised vesicles with  $\alpha$ -syn and VAMP-2. (a) Synaptosomes are pinched-off synapses from neurons which reseal into vesicle-filled spheroids. (b) Zoomed-in drawing of a section of the synaptosome membrane, showing how the membrane-intercalating mCLING dye is trapped during vesicle uptake. The proteins  $\alpha$ -syn and VAMP-2 reside on the vesicle outer membrane, fluorescently labelled with conventional antibodies.

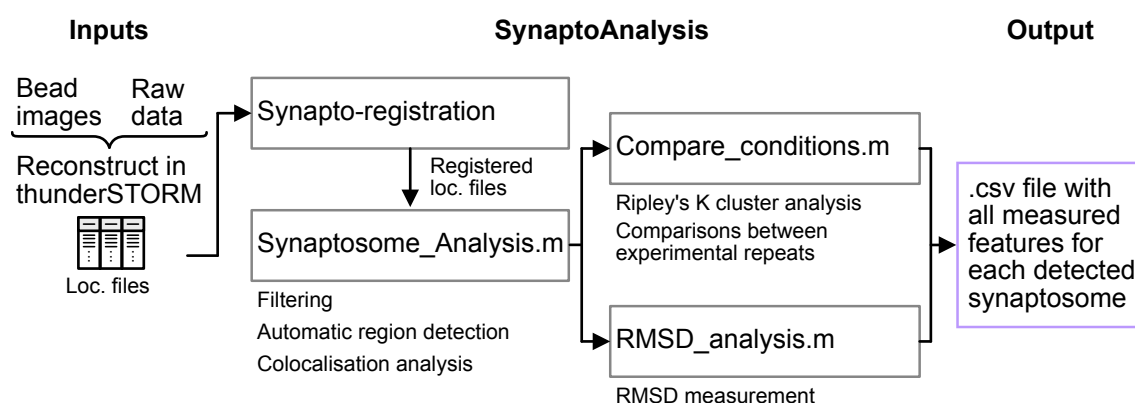
$\alpha$ -Syn and VAMP-2 were stained with antibodies to visualise their spatial distribution in synaptosomes. The membrane-intercalating marker mCLING tagged with an Atto 647N dye [51], was used to label the internalisation of vesicles. mCLING was added to the extracellular media during incubation, exposing the plasma membrane to the dye. When the plasma membrane was internalised during endocytosis, the mCLING dye was also taken up and encapsulated in vesicles. Fig. 3.9 illustrates the labelling for  $\alpha$ -syn and VAMP-2, and how mCLING is used to visualise vesicle uptake in synaptosomes.



### 3.4 Development of SynaptoAnalysis, software tool to analyse single-molecule data from synaptosomes

*d*STORM experiments were performed on the synaptosomes, and the acquired data sets were reconstructed in ThunderSTORM [56]. To process and analyse the resulting localisation data tables, a bespoke software toolkit, SynaptoAnalysis, was developed to extract quantitative information from localisation microscopy data sets. SynaptoAnalysis was written in Matlab; it includes pre-processing routines to remove chromatic aberrations, automatically detect synaptosomes, and filter the localisation data to remove artefacts and false positives. Additionally, SynaptoAnalysis incorporates scripts to characterise the spatial distribution of different fluorescent markers and measure colocalisation and cluster size. For colocalisation analysis, both an overall overlap coefficient and Manders  $M_1$  and  $M_2$  coefficients were used to assess the fractional overlap between fluorophores. For cluster size analysis, a root-mean-square distance from centroid (RMSD) measure was proposed as an alternative to using Ripley's K-function to characterise the spread of spatial point patterns.

An overview of the data processing pipeline using the SynaptoAnalysis scripts is shown as a flowchart in Fig.3.10. Readers mainly interested in the biological results can jump to Section 3.6. For researchers who want to use or further develop the software, this section describes each component of SynaptoAnalysis in detail, with demonstrations of usage and characterisation using test data sets. The source code for SynaptoAnalysis can be found in this repository <https://github.com/pedropabloVR/Synaptosome-Analysis>, where it can be forked and modified by future users.



**Fig. 3.10** Data processing pipeline using SynaptoAnalysis. The localisation files from reconstructed bead images and raw data files in thunderSTORM are used as inputs. The processing and analysis routines (grey rectangles) can be used sequentially, resulting in a .csv file with a list of identified synaptosomes and measured features as output.

The routines included in SynptoAnalysis were developed specifically to solve problems encountered either during data acquisition, or with the reconstructed localisation data in the initial stages of analysis. For example, visible chromatic aberration between the different channels was observed during image acquisition, and therefore a chromatic registration routine was developed as a first pre-processing step after reconstruction. The synaptosomes were also observed to have a high density of localisations across all three channels during *d*STORM acquisitions, therefore a density threshold was applied to remove sparse, background localisations, as well as an intensity threshold using Otsu's method [74]. Upon reconstruction of the data, the first few regions of interest containing synaptosomes were cropped by hand to perform preliminary analyses, however this manual procedure was deemed unfeasible for large data sets. Therefore, an automated synaptosome detection routine was developed which uses the spatial overlap between the three colour channels as a requirement for a cluster to be classified as a synaptosome. These processing routines, which were developed separately over time, were integrated together into SynptoAnalysis to create a streamlined data processing pipeline. These routines were validated either with error functions computed pre- and post-validation (chromatic registration) and percentages of detected candidates as a function of a given threshold (spatial overlap threshold), or empirically by observing the effect of the applied routine on the data and testing different parameters (e.g. filter radius, localisation density per area, intensity threshold, etc.). The colocalisation and cluster analysis routines were included specifically to address the two experimental objects outlined in section 3.1.4, i.e. to measure the colocalisation of a-syn and VAMP-2 with synaptic vesicles, and to measure the cluster size of a-syn and VAMP-2 in physiological and calcium-depleted conditions. Both analysis routines were validated using simulated, ground truth data.

### 3.4.1 Registration of multi-colour single-molecule data

To prepare the multi-colour channel single-molecule data for colocalisation analysis, chromatic offset as described in Chapter 2 Section 2.3.1 had to be minimised using registration. During the acquisition procedure, chromatic offset was minimised by imaging the three channels sequentially, such that all wavelengths traveled the same optical path onto the same detector. This also ensured the three colour channels had similar monochromatic aberrations across the FOV and the residual aberrations were primarily chromatic.

A registration method was implemented based on Annibale et al.'s approach [75], in which a mapping function is determined between the localised coordinates of fluorescent fiducials at the different desired wavelengths, relative to the coordinates from a fiducial imaged at a reference wavelength. The coordinates from the reference channel act as a

ground truth location of the fiducial positions. This mapping function can then be applied to register the experimental data.

Several images from sparse distributions of 100-nm fluorescent beads were captured at the end of each imaging session to serve as fiducials, with the same laser lines used for the biological sample. These images were then localised with high precision using thunderSTORM [56]. To sample the chromatic aberrations over the entire FOV, the coordinates of several images of bead distributions were superposed to simulate dense sampling of the available coordinate space, removing any overlapping coordinates. A function was then used to map the localised coordinates of the warped colour channels (561 and 488 nm) to the reference channel (647 nm) using the local weighted mean algorithm proposed by Goshtasby [76], and first applied by Churchman et al. [77] in the context of single-molecule localisation microscopy. This algorithm uses a set of control points spread uniformly across the FOV to locally determine polynomial functions that map each control point in the warped channel to its corresponding control point in the reference channel. The mapping function, represented in Fig. 3.11 d as a normalised vector field, was determined using the Matlab function *fitgeotrans*, which calculates polynomial transforms between two matrices.

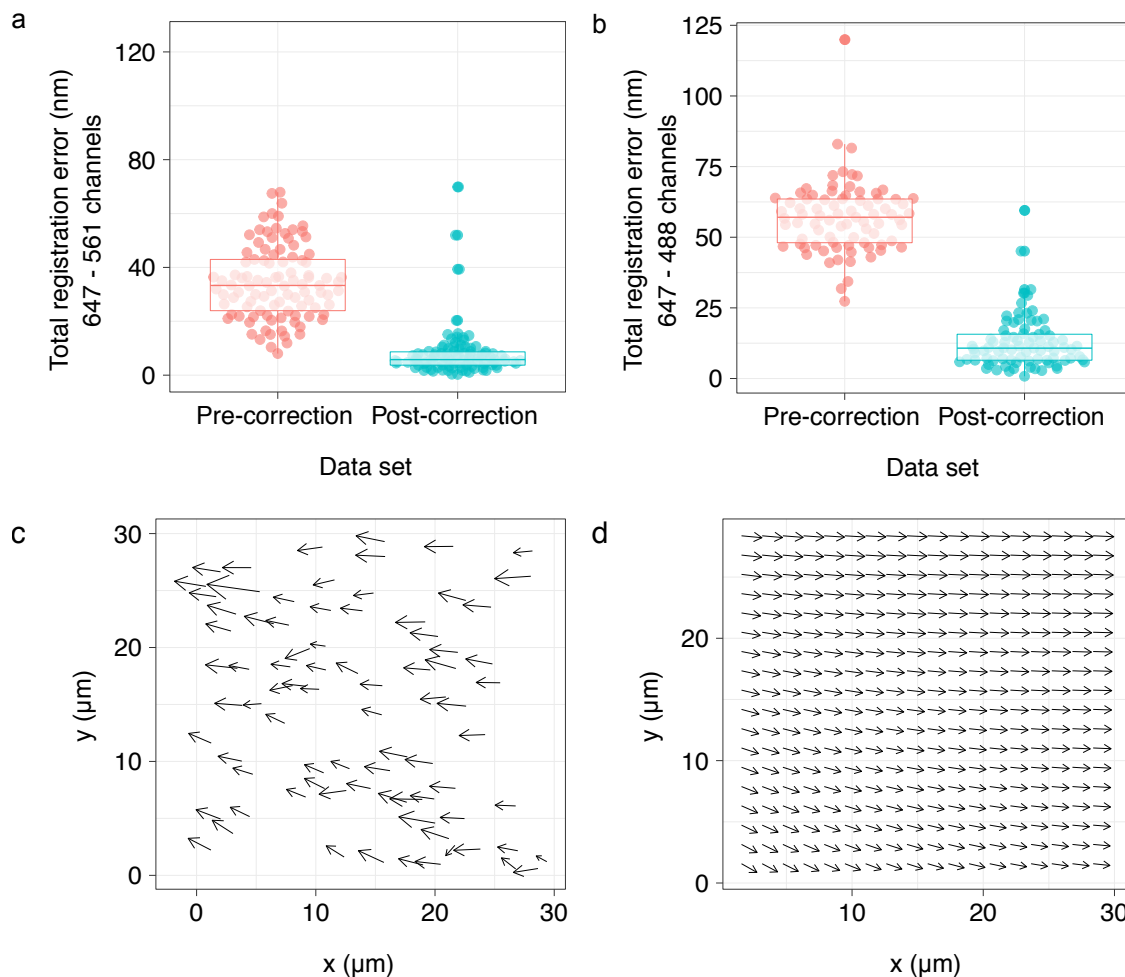
The total registration error (TRE) was implemented as a figure of merit for the mapping function, which is defined prior to registering the reference and the warped channel as follows:

$$TRE_{pre} = \sum_i \sqrt{(x_r^i - x_w^i)^2 + (y_r^i - y_w^i)^2} \quad (3.1)$$

where  $(x_r, y_r)$  are the coordinates for the reference channel control points, and  $(x_w, y_w)$  are the coordinates for the warped channel control points. The TRE post-registration is defined as:

$$TRE_{post} = \sum_i \sqrt{(x_r^i - u^i)^2 + (y_r^i - v^i)^2} \quad (3.2)$$

where  $(u, v)$  are the resulting coordinates after applying the mapping function to the warped channel control point coordinates  $(x_w, y_w)$ . The resulting mapping function was saved for each of the warped channels (561, 488 nm) and then applied to the experimental data sets. The results for a representative calibration are shown in Fig. 3.11, with a mean TRE pre and post correction of 35 and 8 nm for the 561 nm channel, respectively, and 57 and 13 nm for the 488 nm channel. The improvement in the TRE pre- to post-registration is ~4-fold for both the 561 nm and the 488 nm channel. Figures 3.11 c illustrates the raw registration error between the reference and the warped channel in the form of a vector field, and Fig. 3.11 d illustrates the polynomial mapping function used to correct for the shift in Fig. 3.11 c. A number of



**Fig. 3.11** Localisation registration procedure to correct for chromatic aberrations in SMLM data. The total registration error (TRE) in nanometres before correction and after correction was plotted for the (a) 561 nm channel and (b) the 488 nm channel, where each data point represents the TRE for the location of an individual bead. (c) Vector field representation of the raw optical offset between the warped and the reference channel. (d) Normalised vector field representation of a sample correction transform.

outliers are present post-correction in Fig. 3.11 a and b, likely due to interpolation errors in the computation of the polynomial transform at the edge of the field of view between the reference and the warped coordinates.

The registration scripts were written by me and Ezra Bruggeman, using previous code written by Dr. Romain Laine as a starting point and guide. The code is available at <https://github.com/pedropabloVR/SR-registration>.

### 3.4.2 Automatic detection of synaptosomes

After applying chromatic registration, a method was needed to discern between useful regions of interest containing synaptosomes, and background noise from loose dye or random cellular debris in our experimental data. The homogenisation and fractionation steps (Fig. 3.5) during the synaptosome preparation protocol outlined in Section 3.2.1 were prone to including brain tissue debris such as membrane structures of the same density as the synaptosomes [78], making it difficult to distinguish between useful and spurious signals. An automatic synaptosome detection algorithm was developed based on filtering of single-molecule localisation data sets followed by segmentation to detect localisation clusters. These clusters were then classified as synaptosomes or debris based on their size, shape, and localisation content. The filtering steps are described below:

#### Constraints on single-molecule localisation parameters

Spurious localisations with low intensity or high uncertainty can often arise from either background or out-of-focus signal contributions in the sample. SynaptoAnalysis imports the registered localisation files for the three colour channels and filters them based on frame number, intensity in photon units, sigma value of the Gaussian fit, and uncertainty of the centroid localisation to reduce the incidence of spurious localisations. The constraint values for each parameter, summarised in Table 3.2, were either chosen empirically by looking at histograms of the parameter value distributions for a number of samples or following from examples in the literature [79]. The number of localisations removed due to these filters varied widely depending on the density of emitters in the field of view and the amount of background signal from the coverslip. The filters removed anywhere between a few thousand localisations for a FOV with a few synaptosomes, up to hundreds of thousands of localisations for crowded fields of view. Although large, these estimates are not representative of the true useful number of localisations, since most rejected localisations were from background and noise contributions.

**Table 3.2** Filtering values for single-molecule localisation parameters.

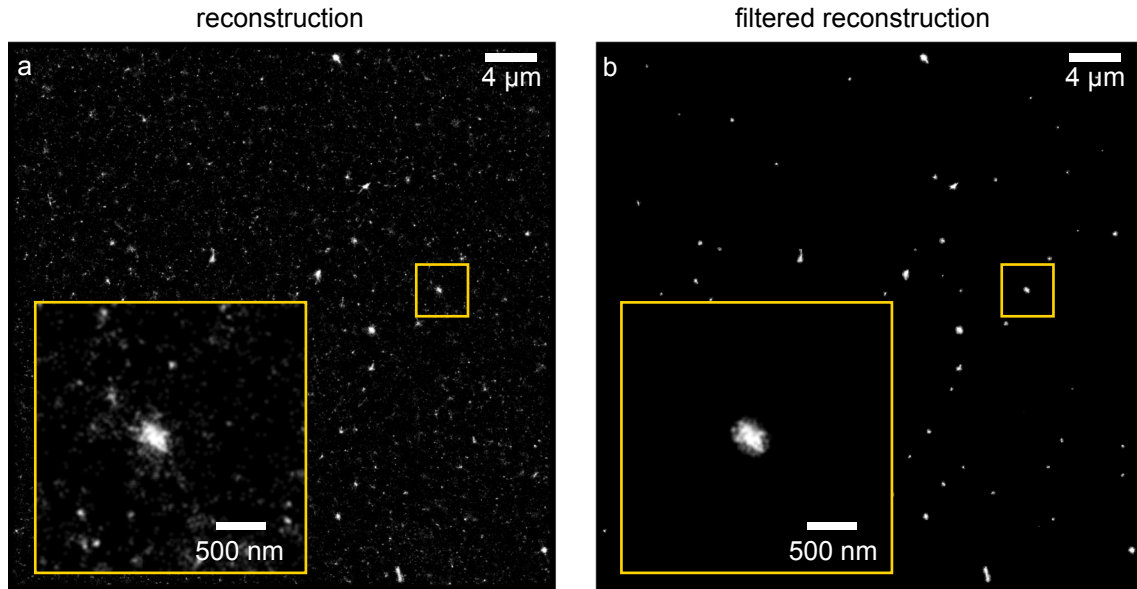
Parameter	Value	Comments
Minimum number of frames	500	Remove first 500 frames to avoid overly dense localisations
Minimum intensity AF568,AF488	500 photons	Min. photons per switching cycle expected from an AF568 and 488 dye [79]
Minimum intensity Atto 647N	1000 photons	Min. photons per switching cycle expected from an Atto 647N dye [79]
Minimum uncertainty	5 nm	Points localised with an uncertainty smaller than 5 nm are likely spurious, or emitters which didn't leave the <i>on</i> state.
Maximum uncertainty	40 nm	The upper limit to our desired localisation precision is the diameter of a synaptic vesicle.

### Density filtering

A density filter based on a nearest-neighbour search was applied to remove sparse localisations in the 647 nm channel (mCLING), which were likely non-specifically bound Atto 647N dye clusters. The Matlab function *rangesearch* was used, which finds all neighbours within a specified distance using Eq. 2.20. Localisations with fewer than  $N$  neighbouring localisations within a specified radius  $R$  were discarded. The search radius  $R$  was chosen from the maximum radius of a synaptosome,  $\sim 1500$  nm [80], and the minimum number of neighbours  $N$  was chosen from the average number of synaptic vesicles inside a synaptosome,  $\sim 300$  [80]. The effect of the two filtering steps described above can be visualised in Fig. 3.12, in which only localisation clusters of a given size and density are preserved and all other localisations are discarded. The filtered localisation data tables were used to generate a reconstructed image for each channel, with each localisation drawn as a point on a black canvas convolved with a Gaussian kernel with width proportional to the localisation precision of the spot.

### Segmentation

Following the localisation parameter and density filtering steps, reconstructed images were segmented to remove candidate clusters too small to be synaptosomes. The filtered reconstructed images for each channel (2D histograms) were binarised using Otsu's method [74]

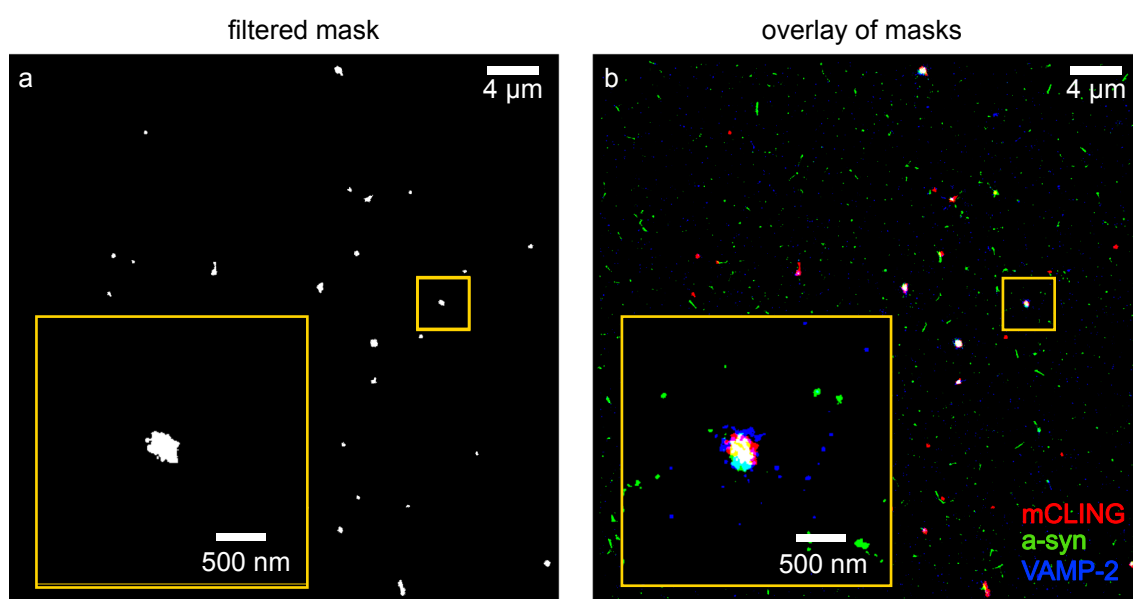


**Fig. 3.12** Effect of filtering SMLM data by constraints on the fit parameters and the density of localisations. (a) Unfiltered 2D histogram representation of a sample data set reconstructed in thunderSTORM. (b) 2D histogram representation after applying constraints on the number of frames, intensity, sigma, and uncertainty of the localised points, as well as a nearest-neighbour density filter to remove sparse localisations.

to create masks, as shown in Fig. 3.13 a. Objects with fewer than 200 connected pixels were removed from the mask of the 647 channel. From the masks created, the Matlab function *regionprops* was used to label the clusters in the binary mask, and to extract the centroid and area for each cluster. No further filtering was applied on the 561 (a-syn) or the 488 (VAMP-2) channels, as we had no a-priori knowledge of the size or distribution of the a-syn or VAMP-2 markers in our samples.

### Colour overlap thresholding

The final step for synaptosome detection was to verify if the candidate clusters contained all the fluorescent markers required for analysis. A minimum amount of overlap between the 647 and the 561 nm channels (mCLING and a-syn) and the 647 and 488 nm channels (mCLING and VAMP-2) was necessary to test for colocalisation. This was accomplished by cropping an 80x80 pixel<sup>2</sup> region around every synaptosome candidate in the binary masks and finding its intersection with the masks of the respective cropped regions in the 561 and the 488 nm channels. The intersection, as described in the Colocalisation section in Chapter 2, is the percentage of spatial overlap between the non-empty regions in the masks. To determine a

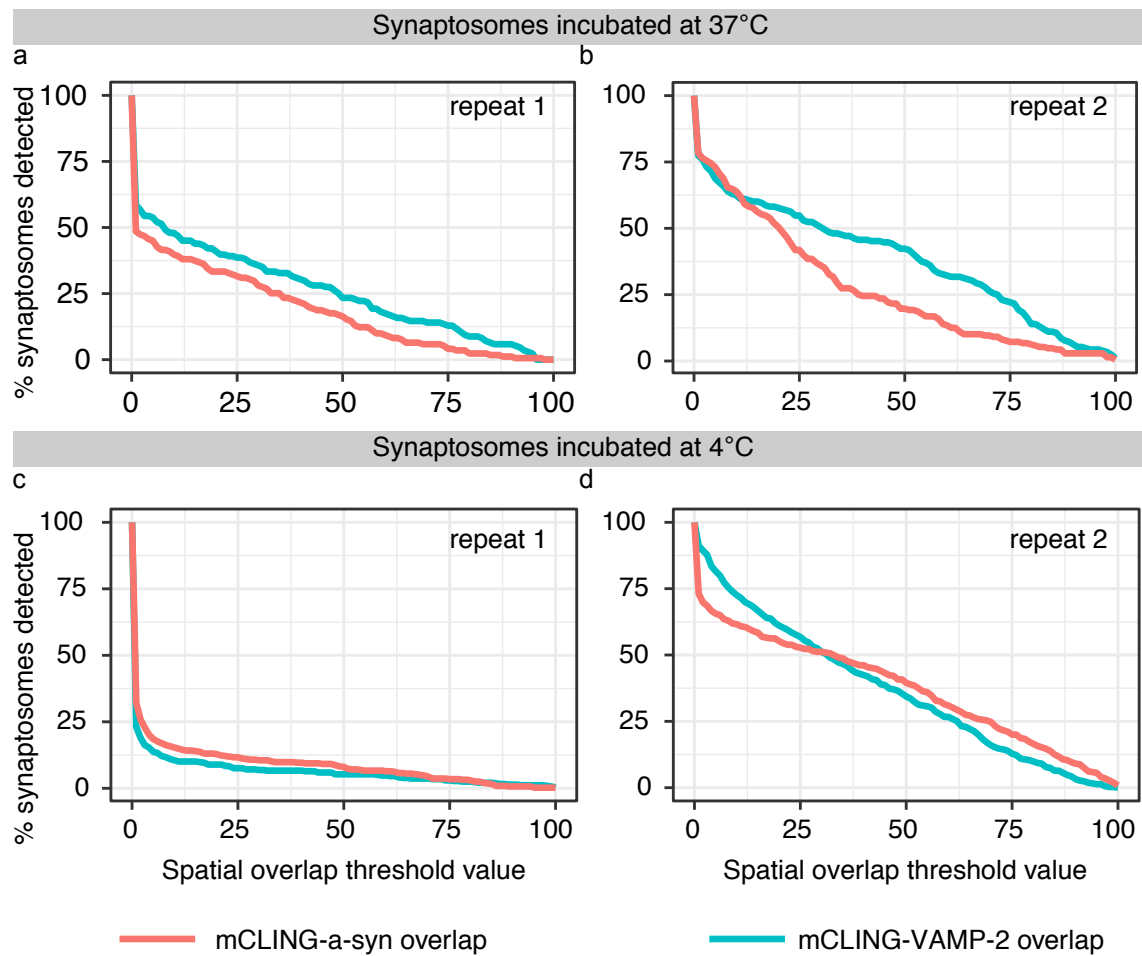


**Fig. 3.13** Segmentation to remove small elements from reconstructed images (a) Representative mask from the 647 nm channel, after removal of elements with fewer than 200 pixels in area. (b) Three-channel overlay of the masks for the 647 (red), 561 (green), and 488 (blue) nm channels.

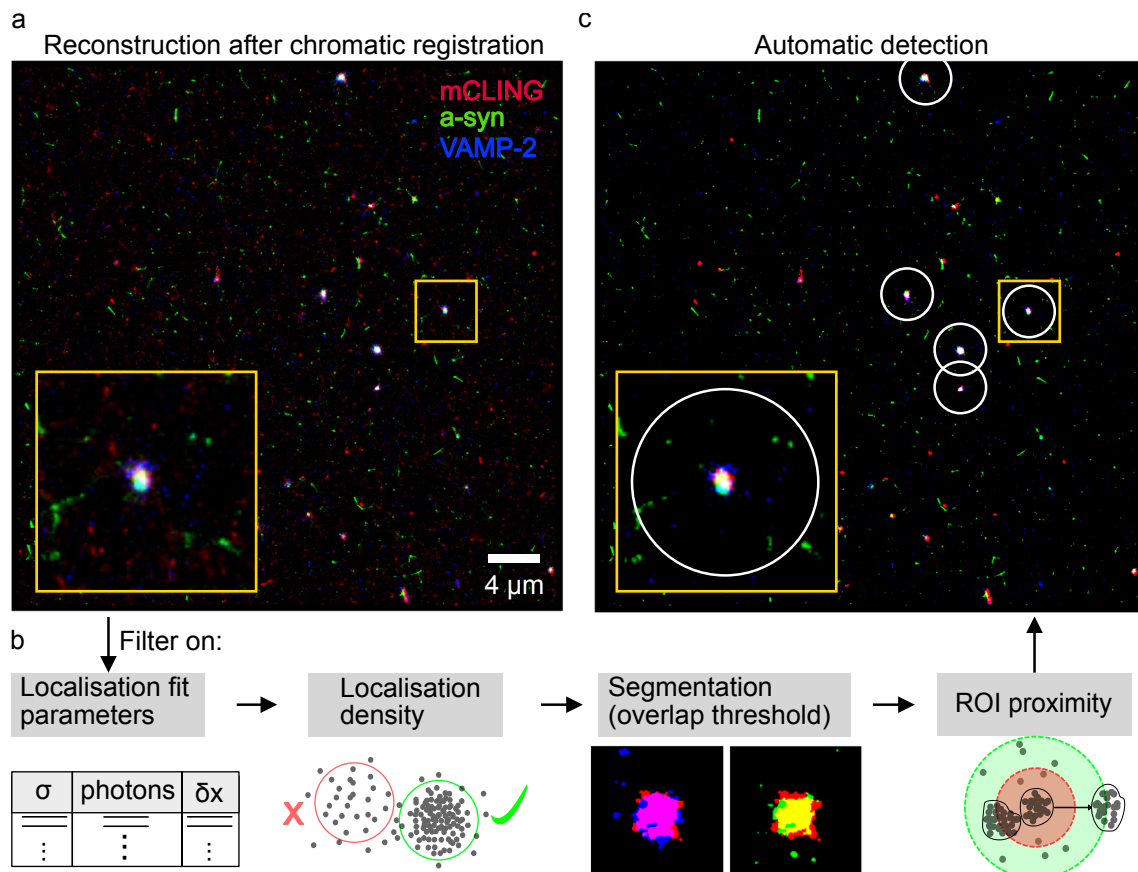
suitable overlap threshold, the percentage of detected synaptosomes was plotted as a function of the spatial overlap threshold for all the data sets collected as shown in Fig. 3.14.

To avoid discarding a large number of potentially useful candidates, the threshold was chosen based on the overlap values (x-axis) for which the rate of change in the percentage of detected synaptosomes started to flatten out in Fig. 3.14 a-d. A 10% minimum area overlap threshold was implemented to select candidates which contained both a-syn and VAMP-2. The localisation precision for each colour channel could also have a significant effect on their spatial overlap, as the dyes exhibit different photon yields in each acquisition and will therefore have different localisation precisions. A low photon yield for any dye will result in a wider spread of its measured single molecule distribution (lower localisation precision), than for a dye with a high photon yield, potentially causing an artificial spatial overlap between the dye distributions. To mitigate the effect of this artificial overlap, localisations with a precision larger than 40 nm for all channels were discarded as specified in 3.2. Finally, a proximity filter was implemented to discard candidates which were closer than 300 nm centre-to-centre from a neighbouring candidate to avoid analysing synaptosome clumps. The coordinates of the  $80 \times 80$  pixel<sup>2</sup> regions corresponding to the selected synaptosomes were then extracted from the localisation data tables. A summary of the detection procedure, from the registered reconstruction, to the filtering steps for automatic detection is illustrated in Fig. 3.15.





**Fig. 3.14** Percentage of detected synaptosomes as a function of the overlap threshold value set for the mCLING/VAMP-2 and mCLING/a-syn channels. Both experimental repeats of the 37°C experiment (a-b) and the 4°C experiment (c-d) were analysed.



**Fig. 3.15** Summary of synaptosome detection steps from localisation data using Synpto-Analysis. (a) Reconstruction of *d*STORM data from a representative FOV. (b) The registered localisations are filtered based on localisation fit parameters and localisation density. A segmentation step followed by a channel overlap threshold, and a nearest-neighbour proximity filter allow us to automatically detect synaptosomes as shown in (c).

Up to this point the SynaptoAnalysis algorithm to detect synaptosomes can be summarised as follows:

1. Superpose localisation files from multiple bead images as fiducials, and compute a geometric transformation to register warped channels (568 and 488 nm) to a reference (647 nm),
2. Compute  $TRE_{pre}$  and  $TRE_{post}$  (Eqs. 3.1 3.2). If  $TRE_{post} < 20$  nm, save transform and apply on experimental data to reduce chromatic offset,
3. Filter localisations by frame number, intensity, Gaussian fit sigma, and localisation precision to remove poor-quality data,
4. Apply a density filter using a nearest-neighbour search (Eq. 2.20) to remove localisations from loose dye particles or tissue debris,
5. Segment reconstructed images using Otsu's method, and remove objects with fewer than 200 connected pixels,
6. Compute the overlap between mCLING/a-syn and mCLING/VAMP-2 channels in all detected synaptosome candidates, and only select those with a minimum 10% channel overlap,
7. Apply a proximity filter using a nearest-neighbour search to remove candidates located within 300 nm of another candidate,
8. Finally, extract coordinates of the  $80 \times 80$  pixel<sup>2</sup> region around each synaptosome candidate, and store individual localisation tables as .csv files for each synaptosome detected.

After removal of spurious localisations and false positives, and extraction of the relevant image regions containing synaptosomes, colocalisation and cluster analysis can now be applied on each individual region. A typical analysis using SynaptoAnalysis could take between 30 and 90 minutes, depending on the number of synaptosomes detected in each field of view. The most computationally intensive step was the density filter, as it iterates over every localisation and counts its nearest neighbours, which can be significantly time-consuming for high-density data sets such as synaptosomes.

### 3.4.3 Colocalisation analysis

A robust metric was needed to quantify the co-occurrence between the spatial distributions of mCLING/a-syn, and mCLING/VAMP-2. Multiple software tools exist to test for colocalisation in SMLM, as described in Table 2.2 of Chapter 2. Using existing colocalisation tools

requires running independent scripts to obtain results, breaking the workflow of our software pipeline. This is reasonable for one-time uses; however, it becomes cumbersome for running analyses multiple times on large datasets such as those in SMLM, e.g. when evaluating the sensitivity of the synaptosome detection on parameters such as the overlap threshold (Fig. 3.14). Therefore, routines were integrated into SynaptoAnalysis to calculate both the localisation and pixel-based co-occurrence between the localised marker distributions. A pixel-based co-occurrence indicator was first chosen to analyse the synaptosome data sets, since we were interested in understanding how the point distributions overlapped, weighted by their relative density. A density-weighted overlap coefficient (WOC) was calculated between the binned localisation distributions of the mCLING and a-syn, and mCLING and VAMP-2, as specified in the equation below:

$$WOC = \frac{\sum_{i=1}^n L_{1,i} \times L_{2,i}}{\sqrt{\sum_{i=1}^n L_{1,i}^2} \times \sqrt{\sum_{i=1}^n L_{2,i}^2}} \quad (3.3)$$

where  $L_1$  and  $L_2$  represent the number of localisations in channel 1 and 2, respectively, for a total of  $n$  pixels analysed. This indicator is the overall overlap coefficient in Eq. 2.25 from Manders et al. [81], as described in Chapter 2. Although the WOC does not provide an absolute quantitative measure of colocalisation [82], it is a useful indicator to compare relative measures of co-occurrence between two markers. To deliver more quantitative results on the fractional co-occurrence between colour channels, Manders' individual  $M_1$  and  $M_2$  coefficients (Eqs.2.21 and 2.23) were also applied on the binned localisations.

### Validating spatial co-occurrence measurement using Manders' coefficients

The colocalisation routines were validated with a sample synaptosome data set in a series of extreme cases. Validation is important, as it ensures the software produces expected results with test data sets. To generate a test data set, the localisation table of a sample FOV was split into even and odd halves to simulate two independent single-molecule measurements of the same underlying structure, as shown in the diagram on Fig. 3.16 a and b. Images (2D histograms) were reconstructed using the even and odd data sets (Fig. 3.16 b), and yellow insets show the point distribution of the same underlying structure, but reconstructed using distinct sets of localisations. This test data set simulated a dual-colour label on the same structure, and allowed us to test the maximal experimentally available colocalisation.

The spatial co-occurrence was measured with the WOC, and Manders'  $M_1$  and  $M_2$  coefficients between the even data set with itself, the odd data set, and the odd data set rotated 90 degrees to simulate a random distribution of localisations following Dunn's method [83]. The even-odd data set demonstrated the maximal experimentally available colocalisation,

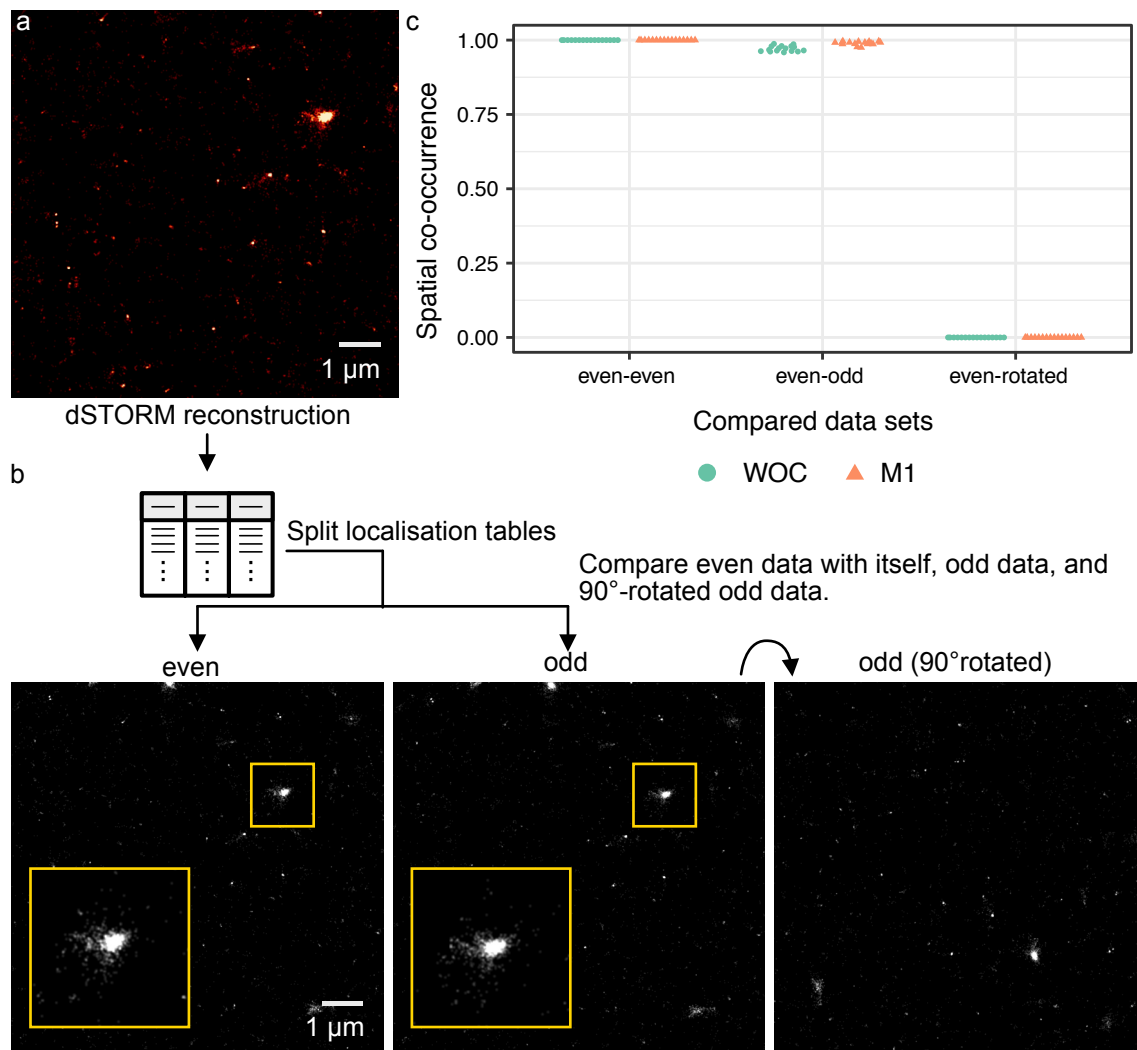
yielding high but not complete colocalisation from two simulated independent measurements of the same underlying structure Fig. 3.16 c. The even-even, and even-rotated odd data sets expectedly yield complete and zero co-occurrence, respectively. This validation using extreme cases in which perfect, high, and zero co-occurrence was expected verified the implemented WOC and Manders' individual  $M_1$  and  $M_2$  coefficients deliver sensible results for comparing the spatial co-occurrence of two point distributions. By effectively splitting the localisation data into two subsets, each with half the localisations, we are reducing the total effective information about the sample and therefore increasing the influence of noise and background contributions to the overall signal. For high density, punctate samples such as the synaptosomes this approach can be satisfactory as demonstrated here, however for continuous structures such as microtubules or actin this approach would be inadequate. An alternative to Dunn's method to test for colocalisation could be to use Costes' method applied to SMLM data, in which pixel regions on the order of the average localisation precision are scrambled and their overlap is tested with an unscrambled image. This approach is robust and could be applied in the future to validate colocalisation for the synaptosome data.

#### 3.4.4 Cluster analysis using the RMSD from centroid

A-syn and VAMP-2 clusters from SMLM data have been previously measured using Ripley's H-function (Eq. 2.28) [5]. Although Ripley's H-function has become widespread to measure cluster size in SMLM [84], it has been reported to overestimate the cluster radius by up to a factor of two [85]. We decided to introduce an alternative metric, the root mean square distance (RMSD) from centroid, to provide an estimate of a cluster's size. The RMSD from centroid of a spatial point pattern can be defined as:

$$RMSD = \frac{1}{n} \sqrt{\sum_{i=1}^n (x_i - x_c)^2 + (y_i - y_c)^2} \quad (3.4)$$

where  $(x_c, y_c)$  are the coordinates of the centroid, and  $(x_i, y_i)$  for  $i \in 1 : n$  are the coordinates of all other  $n$  points in the pattern. The RMSD provides an interpretable measure of cluster size, i.e. the average distance of any point from the centre of the cluster, and unlike Ripley's H-function, does not require an input search radius. The two following sections describe simulations performed to examine the influence of point density and localisation uncertainty on the RMSD measurement.



**Fig. 3.16** Validation test for the colocalisation of two point distributions using the WOC and Manders individual coefficients. (a) A sample synaptosome localisation data set was split into even/odd localisations to simulate independent measurements of the same structure. (b) Reconstructed images for the split localisation tables, with insets displaying the same underlying structure but reconstructed with different subsets of localisations. (c) The spatial co-occurrence is plotted for all synaptosomes identified in the FOV for comparison between the even data set with itself, with the odd data set, and with the odd data set rotated by 90°.

### Simulations using a circle with jittered localisations to determine sensitivity of RMSD

To validate the RMSD as a measure of cluster size, simulations were performed to measure how closely the RMSD approximates the theoretical average distance of any point from the centre of mass of a spatial point pattern. Further, the effects of large localisation uncertainty and high point density on the RMSD were also examined, as these parameters could bias cluster size measurements. Localisation uncertainty was modelled as a random jitter or shift from the true position of a point.

First, a circle was chosen as the simplest shape with a closed-form equation for the average distance from any point inside the shape to its centre, given below:

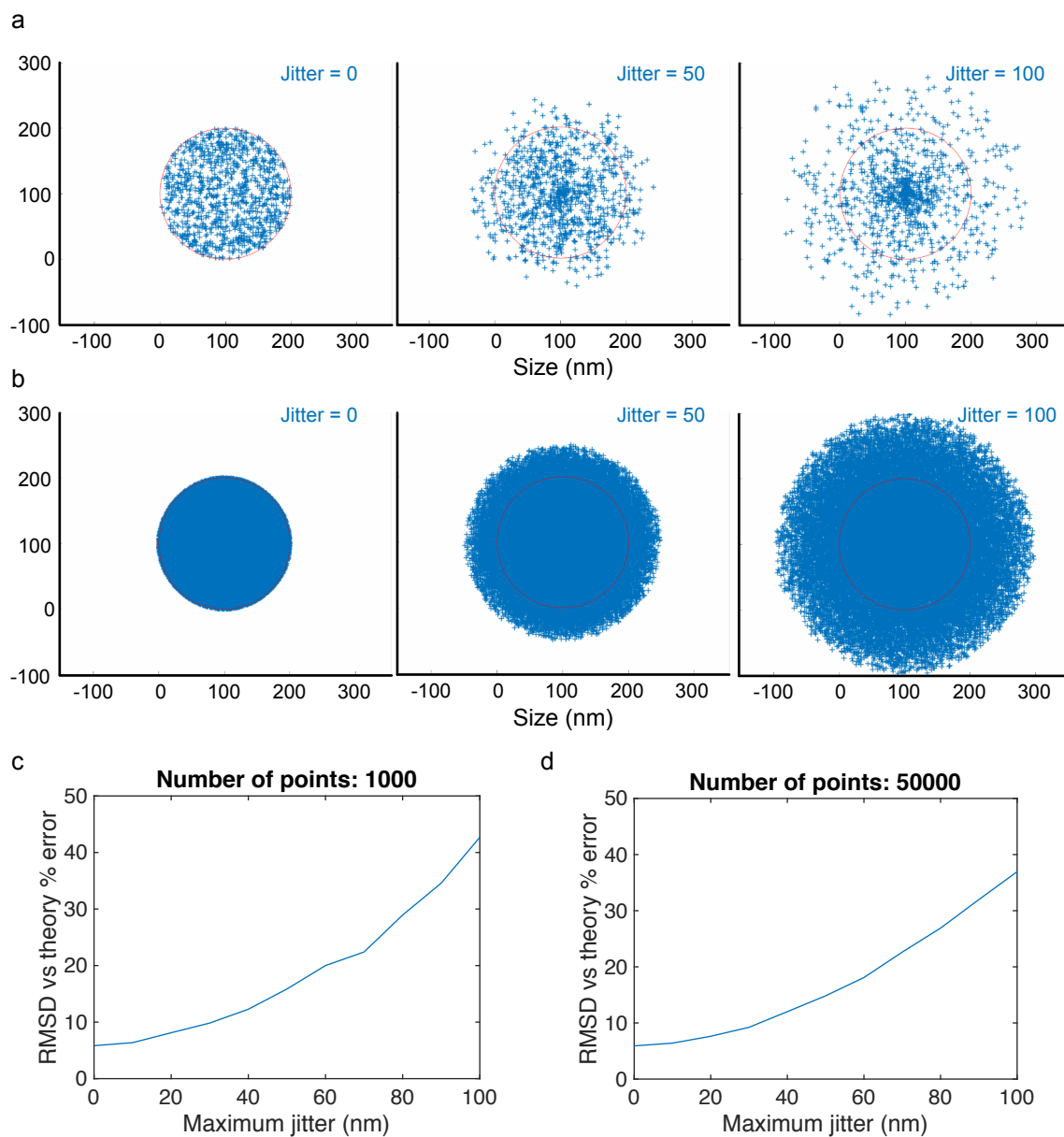
$$\langle R \rangle = \frac{2R}{3} \quad (3.5)$$

where  $R$  is the radius of the circle. A circle with a true radius  $R$  of 100 nm was simulated and a set of  $N$  random points were placed within the boundary of the circle to simulate localisations from a 2D circular object, as shown in Fig. 3.17 a. Each localisation was then jittered in a random direction in 10 nm increments by a maximum amount  $J_{max}$ , with  $J_{max}$  ranging from 0 to  $R$ . The RMSD was calculated during each iteration, and the percentage error between the RMSD and average distance  $\langle R \rangle$  from Equation 3.5 was plotted as a function of the maximum jitter size for  $N=1000$  (Fig. 3.17 a,c), and  $N=50000$  (Fig. 3.17 b,d).

With zero jitter, the RMSD differs by ~5% from  $\langle R \rangle$  in both Fig. 3.17 c and d. As the jitter in the point pattern increases, the discrepancy between the RMSD and  $\langle R \rangle$  increases as the magnitude of the jitter in the point pattern gets closer to the value of the radius of the underlying shape. This is a reasonable result; as the uncertainty in the measurement (the jitter) becomes close to the magnitude of the actual measurement (the radius of the circle), the expected error for the size estimate will be large. In addition to this, the RMSD was not very sensitive to the density of localisations, as the error curve between the calculated RMSD and  $\langle R \rangle$  as a function of jitter was similar for the case a dense point pattern than for a sparse one (Fig. 3.17 c and d).

### Simulations using an arbitrary shape with jittered localisations

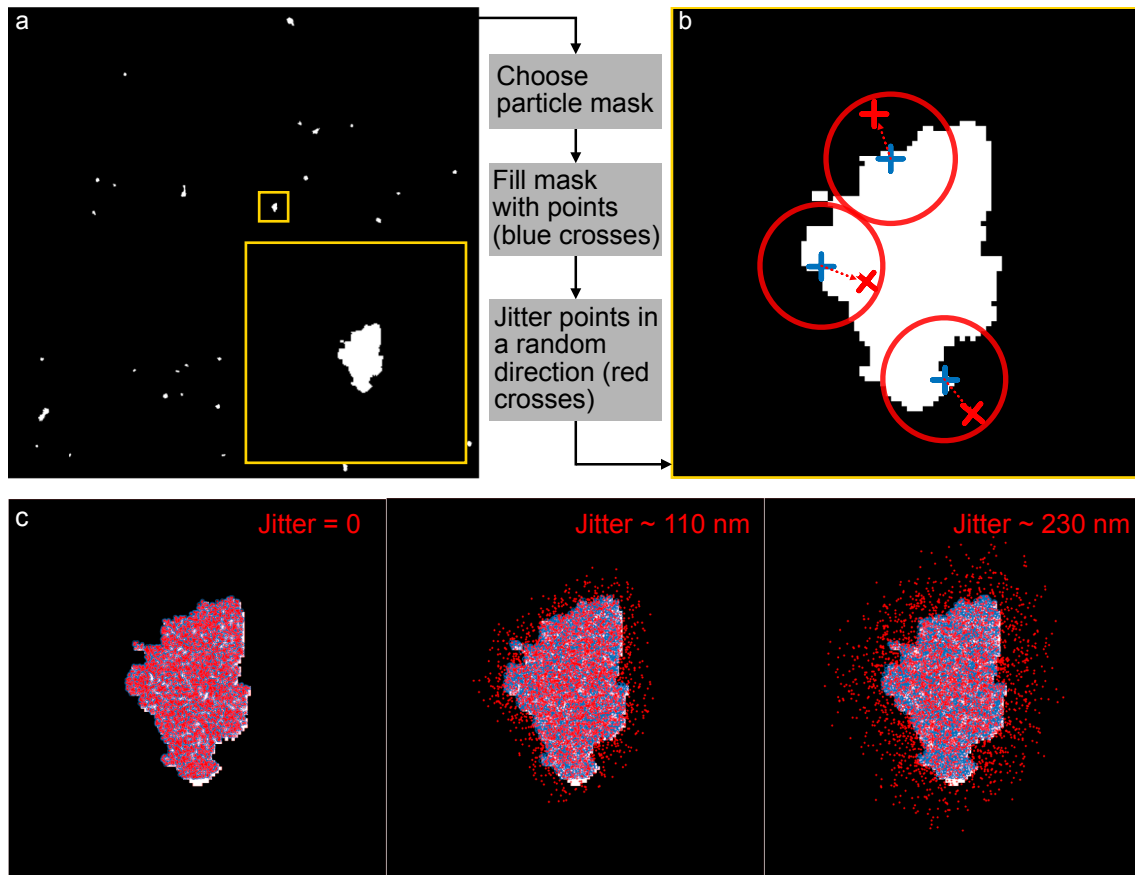
The previous simulation using a circular point pattern was useful to examine how localisation uncertainty and density affected the RMSD. However, localisation distributions from synaptosomes are not perfect circles, but rather arbitrary shapes with inhomogeneous point distributions. Another simulation was performed to evaluate the influence of localisation uncertainty on the RMSD for asymmetric, inhomogeneous point patterns. Instead of simulating synaptosomes, the shapes of real synaptosomes from experimental data were extracted from



**Fig. 3.17** Simulation of a circle with jittered points to calculate the discrepancy between the RMSD metric and  $\langle R \rangle$ . A circle with 100 nm radius was filled with 1000 (a) and 50000 (b) points with a jitter in a random direction increasing from 0 to 100. The percentage difference between the calculated RMSD and the average radius was plotted as a function of maximum jitter introduced, for the case of 1000 (c) and 50000 (d) points.



the segmented mCLING masks in section 3.4.2 (Fig. 3.13 a). Each particle in the binary mask was cropped out in a square window, and artificially filled with points to simulate a spatial point pattern in the shape of the particle (Fig. 3.18 a and b).

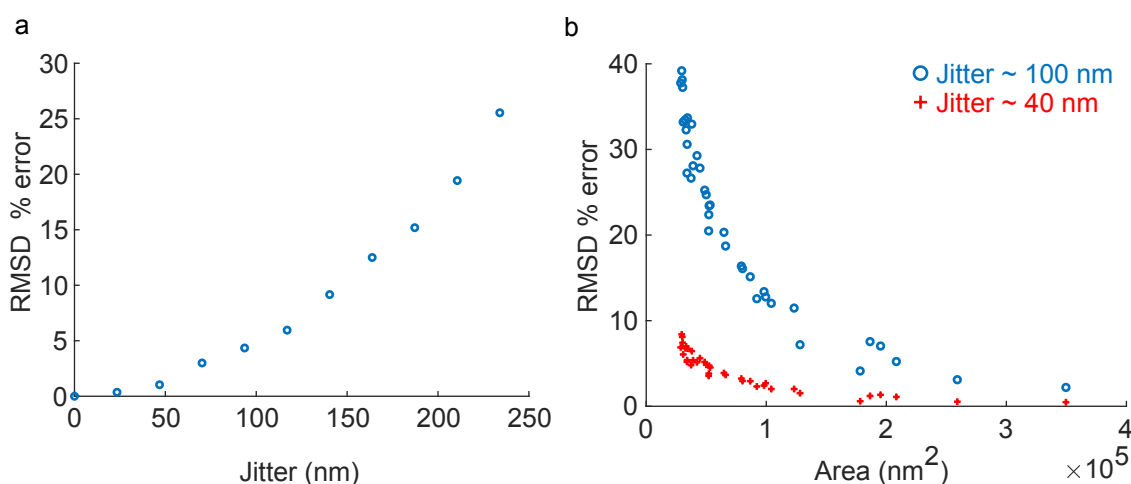


**Fig. 3.18** Simulation using binary masks from synaptosome data to study the influence of localisation uncertainty on the RMSD measurement of cluster size. (a) Binary mask from the mCLING channel with an inset zooming into one of the particles detected. (b) Illustration of how masks were filled sequentially with points within their boundaries, and then the points were jittered in a random direction. (c) Filled-in mask for jitter magnitudes of 0, 110, and 230 nm. The blue dots inside the mask represent the unjittered points, and the red crosses represent the jittered points.

In the absence of a closed-form equation for the average distance from the centre of any point within a randomly-shaped distribution, the RMSD was calculated for the unjittered point patterns in every cluster (blue crosses in Fig. 3.18 b and c) as ground truth estimates of the average distance of any point from their centroid. In a similar fashion to the previous simulation with the circle, the points inside the shapes were jittered in 2-pixel ( $\sim 23$  nm) increments, with a maximum jitter of 20 pixels ( $\sim 230$  nm) in the reconstructed images to simulate localisation uncertainty in SMLM. This jitter was also useful to test extreme cases,

since the largest acceptable localisation uncertainty in the experimental data was only 40 nm (Table 3.2), and a jitter almost 6 times as large was introduced here.

First, the RMSD was measured for all the jittered points in the particle shown in Fig. 3.18 b and c, and the percentage error between theunjittered and the jittered RMSD was plotted as a function of the jitter size (Fig. 3.19 a). The error in the RMSD calculation increases proportional to the jitter, although the error remains small, less than 2%, for a maximum jitter of up to 50 nm. The area of this representative cluster was  $1.83 \times 10^5 \text{ nm}^2$ , roughly corresponding to an equivalent circle radius of  $\sim 250 \text{ nm}$ . From the previous result with the circle simulation, it was clear that the influence of the jitter on the RMSD error depended on the area of the particles analysed. Therefore, the RMSD percentage error was plotted in Fig. 3.19 b as a function of the area in  $\text{nm}^2$  of all the particles found in Fig. 3.18 a, for a maximum allowed jitter of 40 (red crosses) and 110 (blue circles) nm.



**Fig. 3.19** Errors in the RMSD calculation as a function of jitter magnitude, and cluster size. (a) Plot of the RMSD percentage error between the jittered and the unjittered points as a function of the jitter magnitude for the particle shown in Fig. 3.18 c. (b) Plot of the RMSD percentage error between the jittered and the unjittered points for a maximum jitter of 40 nm (red crosses) and 100 nm (blue circles) for all particles in Fig. 3.18 a, as a function of the area of the particles.

The 40 nm jitter magnitude was selected from the upper bound in the localisation uncertainty allowed in our data (Table 3.2). The 110 nm magnitude was used to test and visualise the effect of an extreme jitter on the RMSD calculation. Using a 110 nm max jitter, the RMSD error is large ( $\sim 40\%$ ) for particles with an area smaller than  $1 \times 10^5 \text{ nm}^2$ , which correspond to equivalent circles with radii  $< 170 \text{ nm}$ . For clusters with a larger area, the RMSD error is smaller than 10%. Using a 40 nm max jitter (upper uncertainty bound for experimental data), the RMSD error remains below 10% even for the smallest

particle analysed. This indicates the localisation uncertainty in our experiments is unlikely to influence or bias the RMSD measurement of cluster size.

The two validation simulations performed indicate the RMSD is a robust measure of cluster size, even for small point patterns with a large localisation uncertainty. Additionally, the RMSD was shown to not be very sensitive to point density. Note that the RMSD is not an absolute measure of cluster size, but rather an indicator of the spread of a point pattern that can be useful to compare between different fluorescent distributions. These simulations were performed in Matlab, using the script `part1ch3_validation_SynaptoAnalysis_rmsdSimulations.m` which can be found in the supplementary code repository for this thesis <https://github.com/pedropabloVR/supplementary-code-thesis>.

In summary, a bespoke software package was developed to detect synaptosomes from localisation microscopy experiments, with filters to remove background contributions, and methods to analyse colocalisation and cluster size for multiple colour channels.

### 3.5 Computer model of a synaptosome for generating synthetic SMLM validation data

Concurrent with the development of SynaptoAnalysis and the acquisition of experimental data, a model of a synaptosome stained for mCLING, a-syn, and VAMP-2 was developed *in silico*. Detailed descriptions of the physical features and protein composition of an average synaptosome have previously enabled scientists to build complex 3D synaptosome models [86, 80]. Here, a simplified model was created to generate synthetic SMLM data, which were used to validate the colocalisation and cluster analysis routines included in SynaptoAnalysis. Readers interested in the experimental results of this study can skip this section and go directly to Section 3.6.

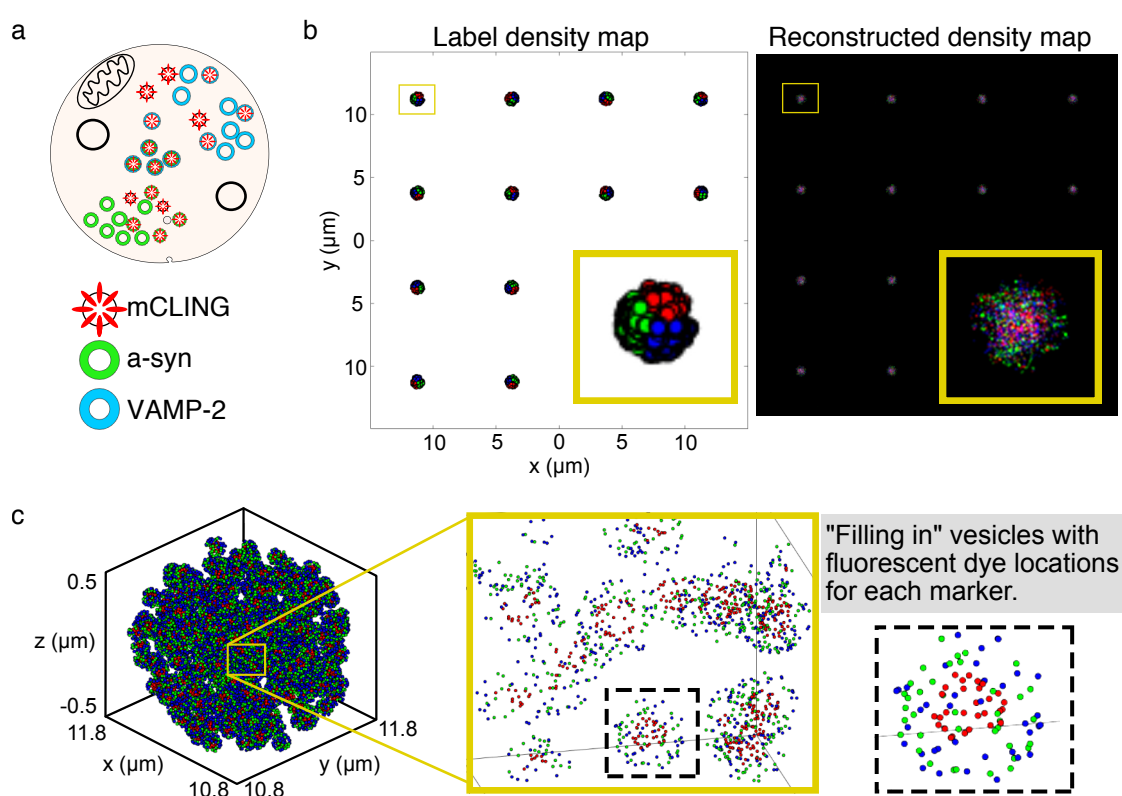
The software package testSTORM [87] was used to build this simulation with accurate parameters on dye photo-physics, noise and background contributions, and the pixelated detector. testSTORM simulates the experimental imaging process in localisation microscopy, generating 3D coordinates for fluorescent molecules and simulating their random attachment to a geometrical structure. The blinking behaviour of each molecule is then simulated using known dye parameters [79], and a stack of diffraction-limited images with single emitters is generated, which can then be reconstructed using localisation software.

This section describes the design of the synaptosome model and the simulations performed using testSTORM to validate the results from SynaptoAnalysis. This work was

performed in collaboration with the developers of testSTORM, Tibor Novak and Dr. Miklos Erdélyi, from the University of Szeged, Hungary.

### 3.5.1 Computer model of a synaptosome

The physical characteristics and protein composition of an average synaptosome and synaptic vesicle have been previously measured by Wilhelm et al. [80] and Takamori et al. [86], respectively, to generate a “library” of features about the synaptic bouton. Based on the parameters measured in these studies, a model of an average synaptosome was created in testSTORM. A 4x4 grid of synaptosomes, each with a 500 nm radius was simulated (Fig. 3.20 b). Each synaptosome was randomly filled with ~380 spherical vesicles with a 20 nm radius. Each vesicle was individually filled with randomly generated coordinates indicating the position of fluorophores, to simulate tagging of mCLING, a-syn, and VAMP-2. This latter “labelling procedure” is illustrated in Fig. 3.20 c, where a zoomed-in density map for a single synaptosome shows the fluorophore positions inside individual vesicles.



**Fig. 3.20** Computer model for a synaptosome. (a) Cartoon of a synaptosome showing vesicles labelled with mCLING, a-syn, and VAMP-2 markers. (b) Equidistant grid of synaptosomes, shown both as a label density map, and as a reconstructed SMLM image. (c) Close up view of a synaptosome filled with vesicles, each in turn filled with the three fluorescent markers.

To keep the model simple, all synaptosomes were assumed to be spherical and of the same size, and to contain the same number of vesicles. Furthermore, all fluorescent emitters in the synaptosomes were assumed to be in focus, and a 1:1 labelling ratio was assumed between the proteins of interest and the dyes simulated (assuming a single binding site per protein). A 2D Gaussian PSF model was used [87].

The red markers for the mCLING were placed close to the centre of the vesicle, since the mCLING intercalates into the inner membrane of vesicles and is short (~2 nm). The green and blue markers for a-syn and VAMP-2 have long linker lengths, resulting from the primary and secondary antibodies used for labelling which are ~10-15 nm in length each along their longest dimension [88]. Therefore, they appear farther away from the vesicle boundary compared to mCLING (Fig. 3.20 c). Previous studies have reported the average number of copies per cell is ~70 for both a-syn [86] and VAMP-2 [89]. However, it is unlikely for antibody-conjugated dyes to reach all 70 copies of both VAMP-2 and a-syn on the surface of a 40 nm diameter vesicle due to steric hindrance. To account for this, and based on the testSTORM constraints on the maximum possible number of labels in a given space, the model only uses 30 copies of a-syn, and 30 copies of VAMP-2 per vesicle. Following the input of the sample parameters, the experimental settings from the microscope (Section 3.2.3) were also fed into the testSTORM simulation to imitate a real experimental acquisition. All parameters used are summarised in Table 3.3.

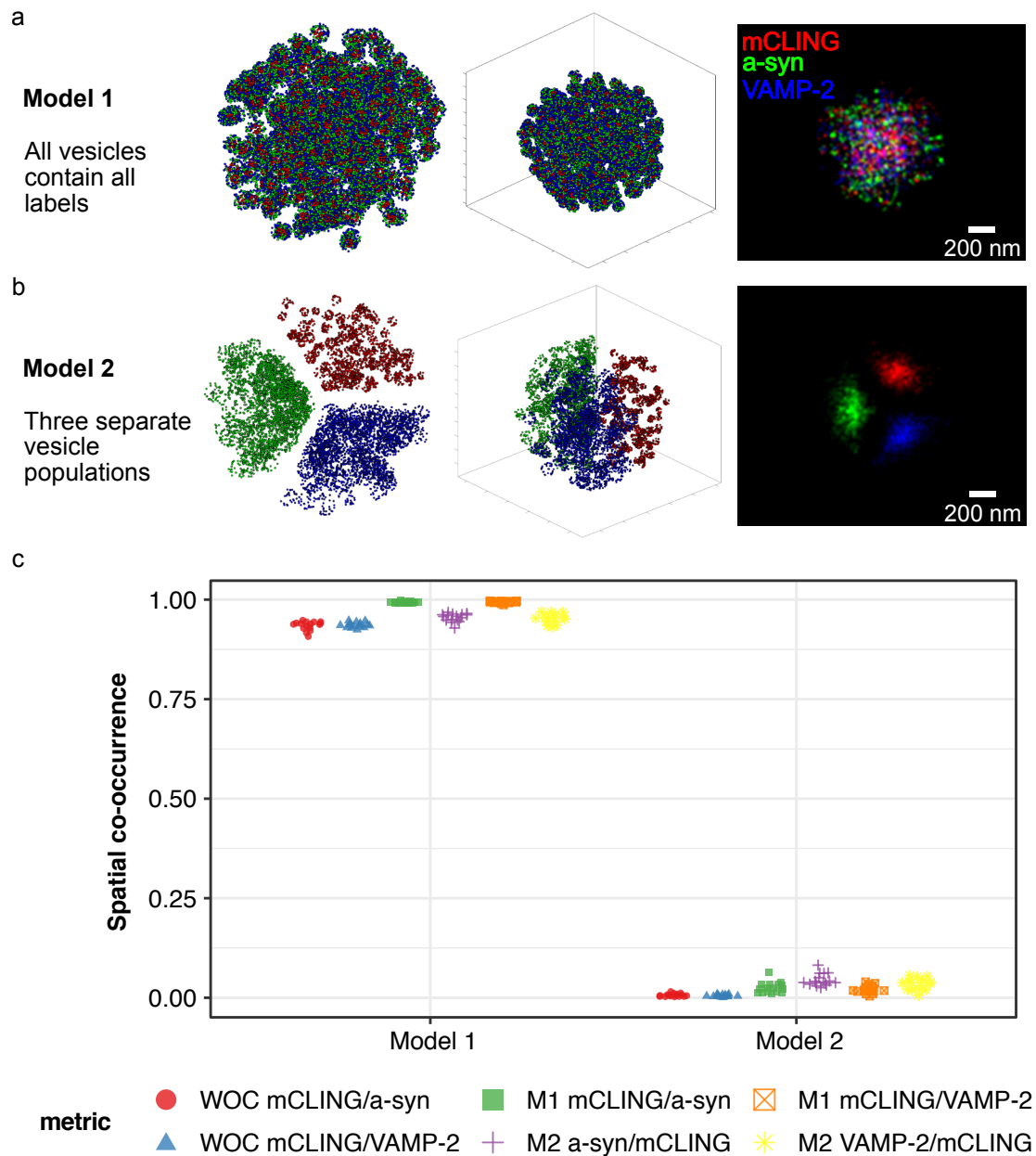
### 3.5.2 Simulating extreme cases of spatial co-occurrence in synthetic synaptosome data

The synaptosome model was used to perform two simulations of extreme cases in the distribution of the mCLING, a-syn, and VAMP-2 markers: model 1 in which all vesicles contain all fluorescent markers and model 2 in which three distinct vesicle populations are labelled separately. The two models were simulated in testSTORM, and the resulting image stacks were reconstructed in thunderSTORM [56] and processed using SynaptoAnalysis. Figure 3.21 a shows the label density map for model 1, along with a composite image of its single-molecule reconstruction. Figure 3.21 b shows model 2, with three separate vesicle populations each containing only a marker for a single fluorescent channel, such that the composite reconstruction shows three spatially separated signals for each channel.

Both models were analysed using the WOC (section 4.c) and Manders'  $M_1$  and  $M_2$  coefficients, and the results were plotted in Fig. 3.21 c. For all metrics, model 1 shows almost complete spatial co-occurrence for both CLING/a-syn and mCLING/VAMP-2, while the results for Model 2 show near zero co-occurrence. These simulations corroborate the

**Table 3.3** Parameters used to simulate SMLM data from an average synaptosome with fluorescent labels on internalised vesicles, a-syn, and VAMP-2.

Sample parameter	Value	Comments
Synaptosome radius	$500 \pm 10$ nm	Wilhelm et al. 2014 [80]
# vesicles/synaptosome	$384 \pm 38$	Wilhelm et al. 2014 [80]
Vesicle radius	20 nm	Takamori et al. 2006 [86]
A-syn copies/vesicle	30	Fahkree et al. 2016 [89]
VAMP-2 copies/vesicle	30	Takamori et al. 2006 [86]
Epitopes/protein	1	Assume 1:1 labelling
mCLINGs/synaptosome	3000	Estimated from image data
Linker length	20 nm	Primary+secondary antibody
<b>Experimental parameter</b>		
Acquisition time	20 ms	Value used in experiments
Number of frames	5000	For quick simulations
Pixel size	117 nm	Measured at the image plane
Numerical aperture	1.49	UAPON100X TIRF objective
EM gain	200	ANDOR performance sheet [57]
Camera baseline	200	ANDOR performance sheet [57]
Quantum efficiency	95%	ANDOR performance sheet [57]
Electrons/count	15.76	ANDOR performance sheet [57]
Ref. index of media	1.518	Immersion oil ref. index



**Fig. 3.21** Testing colocalisation in synthetic synaptosome SMLM data. (a) Model 1 in which all vesicles contain all markers, such that their fluorescent distributions overlap completely. (b) Model 2 in which vesicles are divided into three separate populations, one labelled only for mCLING, one for a-syn, and one for VAMP-2, to simulate a case of zero overlap between channels. A 2D density label map, an isometric view of the density map, and a composite of the reconstructed single-molecule data are shown. (c) Plots of the spatial co-occurrence measured with the WOC and Manders  $M_1$  and  $M_2$  coefficients. Each data point shows an overlap coefficient for a single synaptosome. The script used to generate this plot (part1Ch3\_results\_SynaptoSimulations\_colocalisation.R) can be found in the repository for this thesis <https://github.com/pedropabloVR/supplementary-code-thesis>.

capabilities of SynptoAnalysis to measure extreme conditions in the colocalisation of fluorescent markers in synaptosomes. It is difficult to predict the proportions of fluorescent dyes that will reach each of the three species we are studying. Even though it is well documented that 95% of all vesicles in a synaptosome contain VAMP-2 [86], the amount of a-syn and mCLING in a synaptosome is unknown. This makes it difficult to compare the results from the simulations to the experimental results obtained in real synaptosomes. There are too many steps in the synaptosome preparation, fluorescent labelling, and imaging/reconstruction procedure that can influence the colocalisation results reported from our *d*STORM experiment. It would be unfeasible to attempt to model these, therefore we did not create further models using testSTORM.

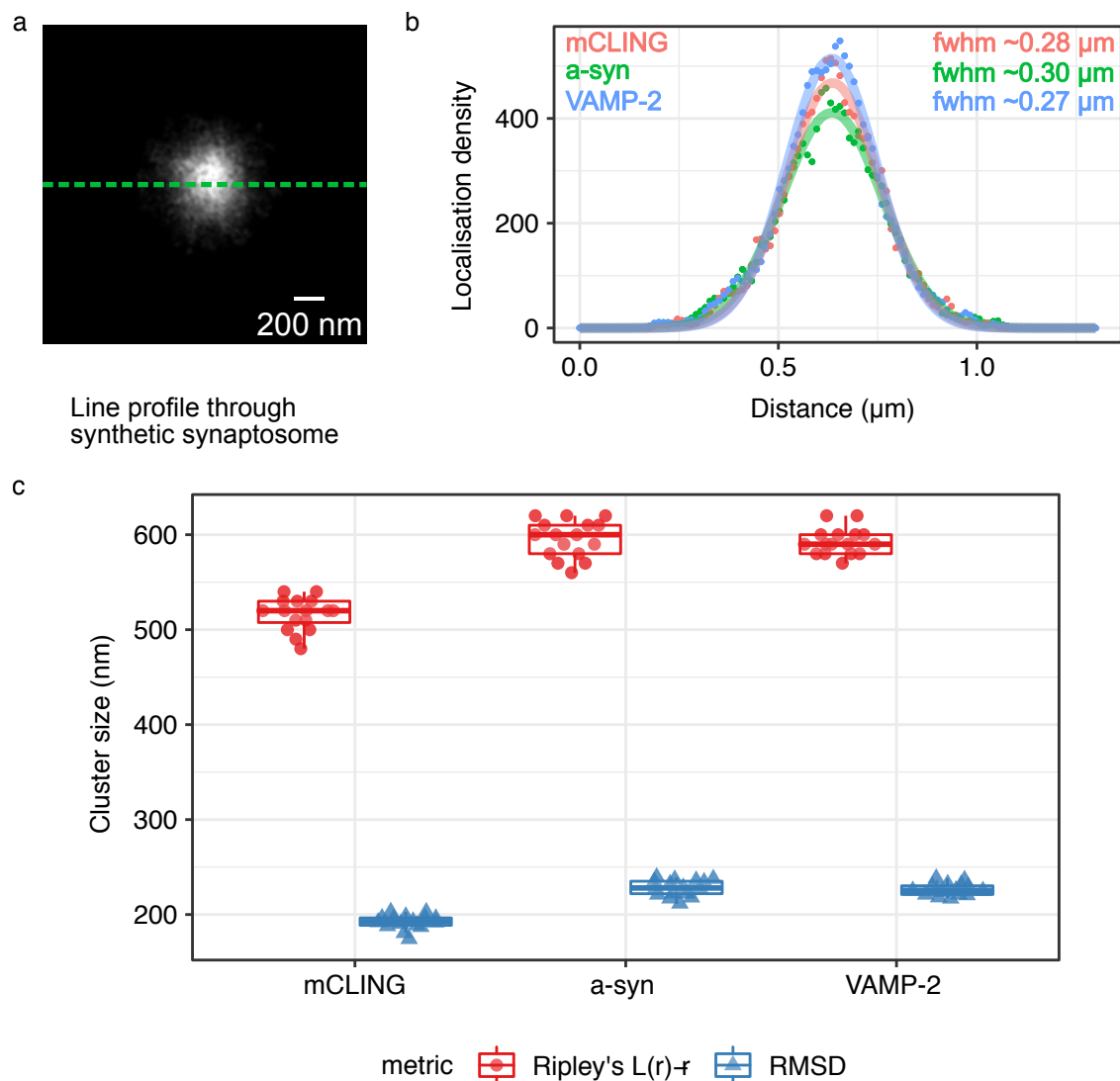
### 3.5.3 Measuring cluster size in synthetic synaptosome data using Ripley's H-function and the RMSD metric

The simulation from model 1 in Fig. 3.21 was used to compare results using the Ripley's H-function and the RMSD metric to measure cluster size differences. Localisation density cross sections through the reconstructed images were fitted to Gaussian functions to provide coarse estimates of each colour channel's distribution size (Fig. 3.22 a & b). The RMSD and Ripley's H-function were subsequently applied to the localisations, and the resulting values for cluster size were plotted in Fig. 3.22 c for each colour channel.

On one hand, the cluster radius estimated from Ripley's H-function for mCLING (mean ~516 nm) was very close to the true radius of  $500 \pm 10$  nm for the synthetic synaptosome. The mCLING has a short linker length (~2 nm) and it resides in the inner membrane of vesicles, therefore its distribution will remain within the synaptosome bounds; this makes mCLING an accurate reference for synaptosome size. The cluster sizes for a-syn (mean ~596 nm) and VAMP-2 (mean ~593 nm) were larger, likely due to the simulated linker length of the dyes (~20 nm) introduced by using primary and secondary antibodies, and the uncertainty in localisation during reconstruction of the synthetic data. On the other hand, the RMSD values for mCLING (mean ~192 nm), a-syn (~228 nm), and VAMP-2 (~226 nm) all significantly underestimated the true synaptosome radius. However, the relative cluster size differences between mCLING and a-syn or VAMP-2 of ~36 nm are closer to the expected difference between these markers due to the extra linker lengths.

In summary, testing Ripley's H-function and the RMSD on a ground truth synaptosome model to measure cluster size demonstrated Ripley's H-function provides a closer estimate to the true radius of a cluster than the RMSD, although the RMSD is sensitive to differences in dye locations due to the long linker lengths used.

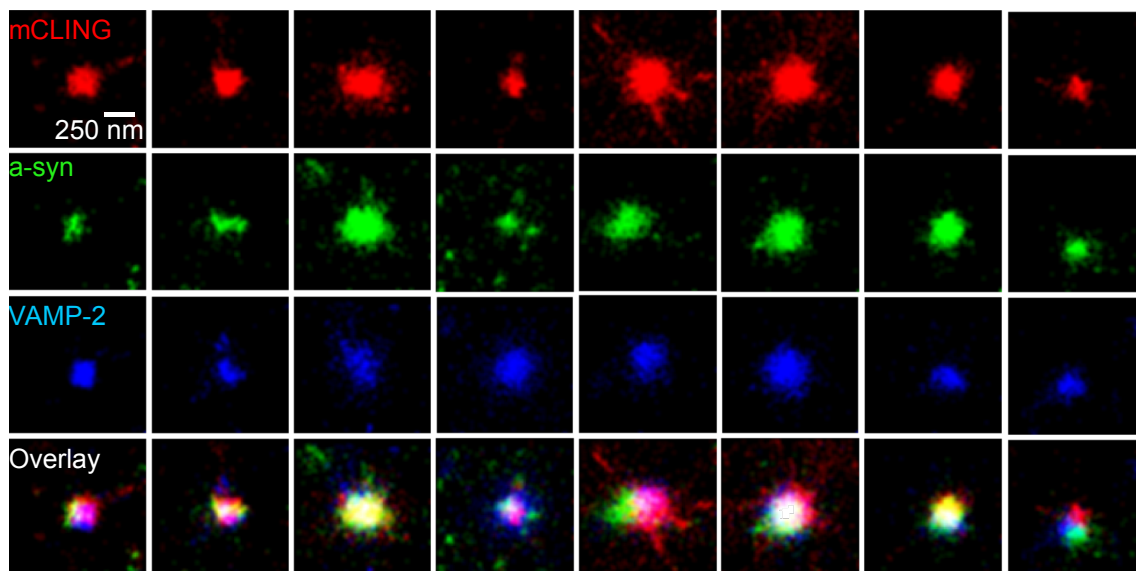




**Fig. 3.22** Simulated cluster size measurements on synthetic synaptosome data from model 1 in Fig. 3.21 a. (a) Representative reconstructed image from a synthetic model 1 synaptosome, with a line profile (green line) taken through its center. (b) Line profiles for each of the three colour channels from the synaptosome in panel (a) were fitted to Gaussian functions and plotted. (c) Scatter and box plots for the cluster size values measured using both Ripley's H-function and the RMSD, for the mCLING, a-syn, and VAMP-2 channels of 16 simulated synaptosomes. Each data point represents a cluster size measurement for a single simulated synaptosome. The script used to generate these plots (part1Ch3\_results\_SynaptoSimulations\_clusterSize\_GaussianFit.R) can be found in the repository for this thesis <https://github.com/pedropabloVR/supplementary-code-thesis>.

## 3.6 Experimental results

The use of *d*STORM and the development of a specialised software tool were driven by the two biological questions set out in Section 3.1.4: (1) How does the colocalisation of a-syn and internalised vesicles compare to that of VAMP-2? (2) How is the clustering of a-syn and VAMP-2 affected by calcium levels at the synapse? This section describes the results obtained by using *d*STORM to image a-syn, VAMP-2, and mCLING in synaptosomes (Fig. 3.23), and applying SynptoAnalysis on the reconstructed data sets. The colocalisation between mCLING/a-syn and mCLING/VAMP-2 was measured at different temperatures to visualise uptake by different recycling mechanisms, and the clustering behaviour for a-syn and VAMP-2 was analysed in the presence and absence of calcium. The findings from these two analyses are presented, which partially answer the research questions stated above.

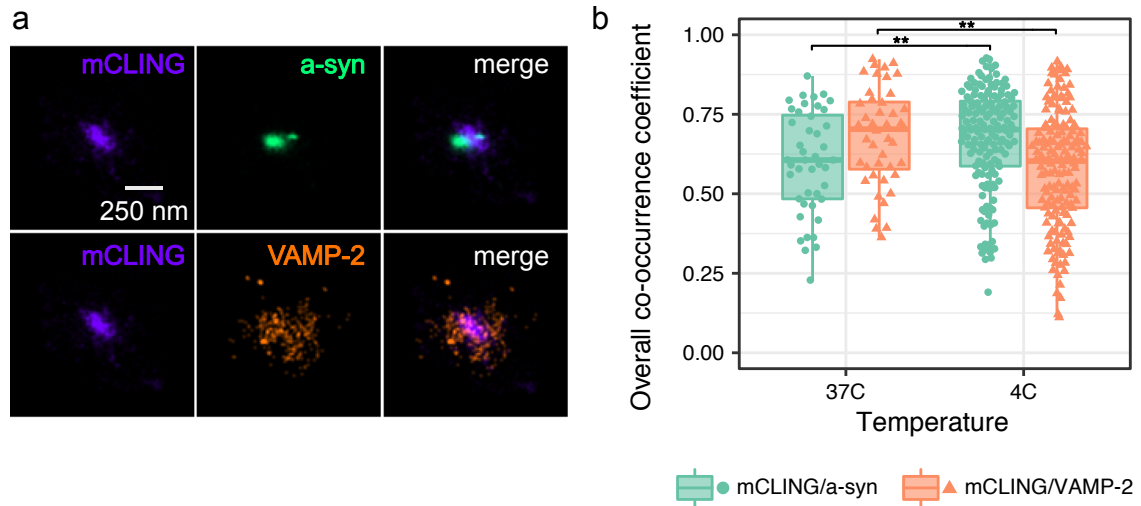


**Fig. 3.23** Reconstructed images of synaptosomes detected with SynptoAnalysis. Single-molecule distributions are shown for the mCLING, a-syn, and VAMP-2 colour channels, as well as an overlay of all three in the bottom row.

### 3.6.1 A-syn's co-occurrence with internalised vesicles is temperature independent, compared to that of VAMP-2

Vesicle homeostasis at the synapse is maintained by multiple membrane recycling mechanisms, which are likely to be active at the same time and to act in collaboration. It is often difficult to identify which recycling mechanisms are active, because dominant CME and SNARE (exocytosis) processes at physiological temperatures can conceal other mechanisms.

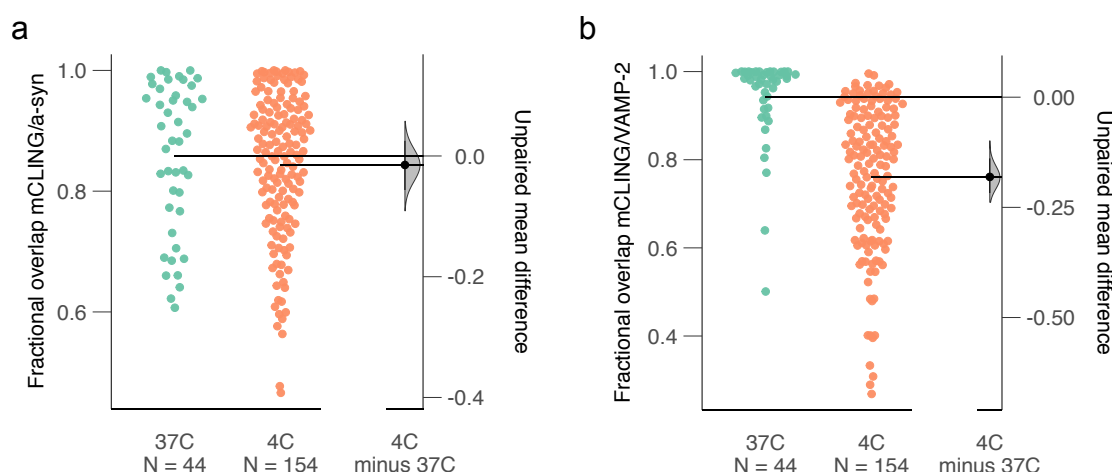
To visualise uptake via different mechanisms, synaptosomes were incubated at 37°C to visualise canonical CME events, and at 4°C, temperature at which membranes become more rigid and both CME and full fusion are inhibited [25], to visualise membrane uptake events from non-canonical mechanisms.



**Fig. 3.24** Colocalisation between mCLING/a-syn, and mCLING/VAMP-2 in synaptosomes using the weighted overlap coefficient (WOC). (a) Representative SMLM reconstructions of the mCLING/a-syn, and mCLING/VAMP-2 channels. (b) Plot of the WOC for two temperature conditions. Each data point represents the WOC for a single synaptosome.

For each detected synaptosome (Fig. 3.24 a), the weighted overlap coefficient (WOF) (Eq. 3.3) was calculated between mCLING/a-syn, and mCLING/VAMP-2, as shown in Fig. 3.24 b. On one hand, the WOC indicated the co-occurrence of mCLING-stained vesicles and VAMP-2 is higher at 37°C than at 4°C. On the other hand, the opposite trend was observed for mCLING/a-syn, with co-occurrence increasing at 4°C. This result provided a useful first indication that indeed, a-syn and VAMP-2 colocalised differently with endocytosed vesicles. However, the WOC (Eq. 3.3) is only a relative indicator of colocalisation, as it provides a number between 0 and 1 which indicates a relative degree of overlap between the channels, without distinguishing the overlap of channel 1 relative to channel 2, and vice versa. Therefore, the WOC is not an absolute metric for colocalisation and is only useful as a relative indicator as specified by Cordelieres and Bolte [82]. To perform a more quantitative colocalisation measurement, Manders' individual coefficients were measured for each synaptosome.

The co-occurrence fraction (Manders  $M_1$ ) between mCLING and a-syn (Fig. 3.25 a) shows no significant difference between the two temperature conditions, whereas for mCLING and VAMP-2 (Fig. 3.25 b) the co-occurrence fraction significantly decreased at



**Fig. 3.25** Colocalisation between mCLING/a-syn, and mCLING/VAMP-2 in synaptosomes using Manders' individual coefficients. (a) Manders'  $M_1$  coefficient for the fractional overlap of mCLING on a-syn, (b) Manders'  $M_1$  coefficient for the fractional overlap of mCLING on VAMP2. Gardner-Altman estimation plots display the measured colocalisation in a scatter plot, and the effect size with its 95% confidence interval between temperature conditions is displayed as a point estimate and vertical bar, on a separate axis. Each data point represents the Manders coefficient for a single synaptosome.

4°C compared to that at 37°C. The high co-occurrence fraction between internalised mCLING vesicles and VAMP-2 at 37°C could be explained by VAMP-2's necessary uptake during CME, so that it can be later used for SNARE exocytosis. The subsequent decrease at 4°C is likely due to a reduction in the number of CME/SNARE events due to decreased membrane fluidity and lower propensity at low temperatures of CME and full fusion exocytosis [25]. The temperature insensitive co-occurrence between mCLING and a-syn may be due to its involvement in non-canonical membrane uptake mechanisms which can be observed even at sub-physiological temperatures, regardless of membrane fluidity [25].

### 3.6.2 A-syn and VAMP-2 clusters increase in size in calcium-depleted conditions

Calcium has been reported to mediate the lipid binding properties of a-syn at its C-terminus, and to influence the clustering behaviour of a-syn at the presynapse [5]. To investigate the effect of differing calcium levels on a-syn, compared to that on VAMP-2, synaptosomes were incubated with physiological and calcium-depleted (EGTA) buffers.

The cluster sizes for a-syn and VAMP-2 were measured using both Ripley's H-function and the RMSD metric in SynptoAnalysis, and the results were visualised using Gardner-

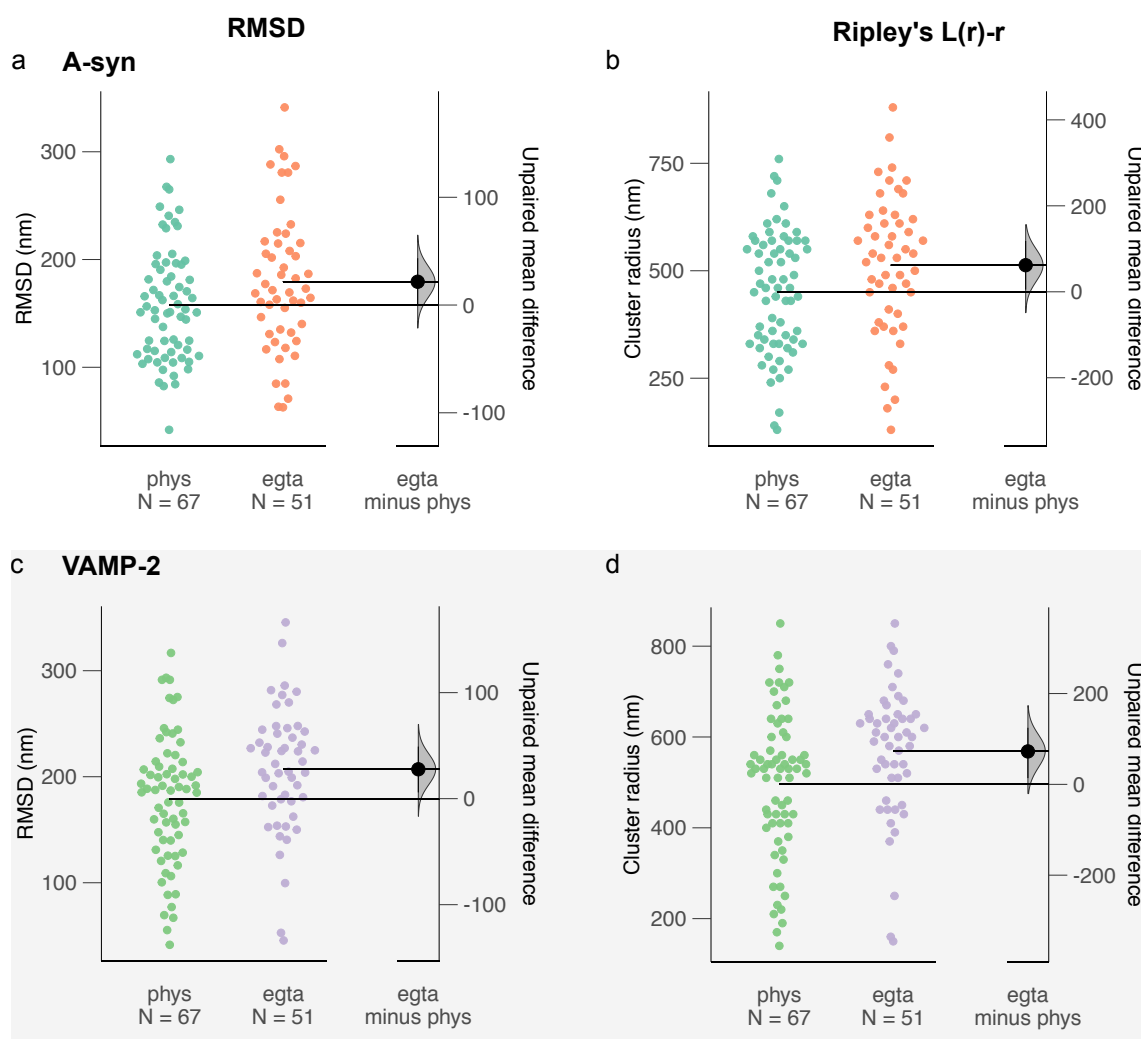
Altman plots in Fig. 3.26 a (a-syn) and c (VAMP-2), with the unpaired mean difference (m.d.), or effect size between conditions shown in a separate axis. For Ripley's H-function, a search radius of 1  $\mu\text{m}$  was applied, this being twice the expected radius for an average synaptosome. The peak of the L(r)-r curve was used to estimate the domain radius [85, 90], and the resulting values were plotted as a function of calcium condition in Fig. 3.26 b (a-syn) and d (VAMP-2). In both a-syn (m.d. ~62 nm) and VAMP-2 (m.d. ~73 nm) an increase in cluster size was observed when calcium was depleted from the incubation buffer. VAMP-2 clusters measured with Ripley's H-function were larger (mean phys ~496 nm) on average than a-syn clusters (mean phys ~451 nm), likely due to VAMP-2 being present in a larger proportion of vesicles than a-syn [86]. Using the RMSD from centroid, an increase in cluster size for both a-syn and VAMP-2 was also observed in the calcium-depleted condition. Again, the effect size was slightly more pronounced for VAMP-2 (m.d. ~28 nm) than for a-syn (m.d. ~22 nm) between conditions. Similarly to the measurement with Ripley's H function, the VAMP-2 clusters (mean phys ~180 nm) were measured to be larger than the a-syn clusters (mean phys ~158).

Note that these two cluster analysis tools measure different properties of the localisation data: Ripley's H-function provides an estimate of the radius of maximal aggregation [85], whereas the RMSD is a measure of the average distance of any point in a pattern from the centre of the pattern. Therefore, the reported values for cluster size were not expected to match between the two metrics, as demonstrated by the validation data in Section 3.5.3. Ripley's H-function once again provided a cluster size estimate closer to the average expected radius of a synaptosome (~500 nm, [86]). Nonetheless, both metrics showed an increase in cluster size in both a-syn and VAMP-2 when calcium was depleted from the incubation buffer.

## 3.7 Summary and Discussion

A-syn is a hallmark of Parkinson's Disease, as it aggregates to form part of characteristic Lewy bodies and neurites found in patients with advanced forms of this disease. Despite its link to Parkinson's pathogenesis, a-syn's function remains unclear.

We set out to investigate if a-syn colocalised with internalised vesicles at the synapse via CME or the SNARE complex like VAMP-2, or if it colocalised with vesicles internalised via different (non-canonical) mechanisms. Furthermore, we aimed to test the influence of calcium on the distributions of a-syn/VAMP-2 positive vesicles. Synaptic vesicles are much smaller than the diffraction-limit, so conventional imaging and analysis tools were not sufficient to explore the nanoscale distributions of the markers of interest. To address



**Fig. 3.26** Cluster size measurements for a-syn (a-b) and VAMP-2 (c-d) localisation data in synaptosomes, using the RMSD (a,c) and Ripley's L(r)-r function (b,d) in physiological and calcium depleted conditions. Calcium was depleted using the chelator EGTA. The number of synaptosomes analysed for each calcium condition is shown on the y-axis. 4 repeats were imaged and analysed for each condition. Each data point represents the cluster size measurement for a single synaptosome.

this imaging challenge, imaging experiments implementing *d*STORM were performed and optimised, and a complete SMLM analysis package, SynaptoAnalysis, was developed and characterised to quantitatively examine our imaging data.

Our results show a-syn colocalises with internalised vesicles independent of temperature, contrary to what was observed for VAMP-2. This suggests a-syn could be associated with vesicle recycling mechanisms which occur at sub-physiological temperatures. In addition, both a-syn and VAMP-2 clusters were found to increase in size in the absence of calcium, which could be due to an interaction between the two proteins. This new evidence advances our understanding of a-syn's role at the presynapse, however further studies and experiments are necessary to determine the recycling mechanisms a-syn is involved in, and to clarify its relationship to VAMP-2. Incidentally, the simulations performed to benchmark SynaptoAnalysis proved useful to gauge the performance of cluster analysis algorithms. The following sections provide a more detailed discussion on the results obtained, with suggestions for future work.

### **3.7.1 A-syn and VAMP-2 associate with different vesicle recycling mechanisms**

At the synapse, multiple pathways may be active simultaneously, and acting in concert, to maintain vesicle homeostasis. It is difficult to observe the role of a-syn at the presynapse because dominant SNARE (exocytosis) and CME processes at physiological temperatures can mask other recycling mechanisms which don't contribute to vesicle uptake/release to the same extent but may also be important.

A high spatial co-occurrence between VAMP-2 and mCLING was observed at 37°C (Fig. 3.25 b), which could be explained by the binding of VAMP-2 to the adaptor protein AP180 during CME, a dominant recycling process at 37°C. This binding ensures VAMP-2 is taken back up into vesicles ready for refilling with neurotransmitters and subsequent rerelease, as VAMP-2 is crucial for SNARE-mediated exocytosis [91]. To investigate other mechanisms which are not dependent on membrane fluidity, the synaptosome incubation temperature was lowered to 4°C and a significant decrease in the co-occurrence of VAMP-2 and mCLING was observed (Fig. 3.25 b).

The colocalisation of a-syn and mCLING, on the other hand, did not change between the 37 and 4°C conditions (Fig. 3.25 a). A-syn has been previously implicated in several steps of vesicle recycling mechanisms [41]; this work shows a-syn colocalises with vesicles recycled via non-temperature dependent mechanisms. The lack of sensitivity to low temperature in the colocalisation between a-syn and internalised vesicles suggests a-syn might be associated with

mechanisms such as kiss-and-run, which are the only known to occur at sub-physiological temperatures [25]. A recent study found overexpressed and endogenous a-syn promotes dilation of the fusion pore in adrenal chromaffin cells and neurons [43], which shows preliminary evidence for a-syn's potential role in kiss-and-run. A-syn's involvement in vesicle recycling remains unknown, however several speculative models could be proposed to argue for its involvement in kiss-and-run. A-syn could potentially tether a vesicle to the plasma membrane via its double-anchor mechanism (Fig. 3.3) to allow the vesicle to open a fusion pore to release or uptake cargo and reseal the pore. Alternatively, a-syn could also act as a restraining tether between two vesicles (3.3), preventing full fusion and collapse prior to delivering neurotransmitter cargo, as previously proposed by Alabi and Tsien [1].

### Future work

Our results suggest a-syn's relation to vesicle recycling mechanisms which are more likely to be observed at sub-physiological temperatures, but the identity of the mechanism remains a mystery. To lend weight to our preliminary results and really understand the relation between a-syn and vesicle recycling mechanisms, further imaging experiments must be performed. To further explore the location of a-syn clusters relative to fusion pore sites, a two-colour 3D imaging experiment could be performed using stimulated emission depletion STED microscopy to observe the formation of large fusion pores (~400 nm) in chromaffin cells, similar to the method reported by Shin et al. [92]. In this way, a-syn's position with respect to the pore could be explored at different stages of pore formation, to elucidate its potential role as a mediator for recycling events. Alternatively, the resolution improvement could be achieved using expansion microscopy [93], a technique in which the sample is artificially enlarged using a polymer matrix to separate fluorophores otherwise too close to each other to be resolved, instead of STED. Fusion pores in chromaffin cells could then be imaged using a traditional confocal microscope with up to a 4x [93] or 16x [94] resolution improvement. To explore the dynamics of the relationship between a-syn and vesicle trafficking, a fast super-resolution technique such as structured illumination microscopy (SIM) could be implemented to visualise in up to three colours the location of a-syn relative to VAMP-2 and endocytosed vesicles at the synapse in live cells. This would require a stable cell line with large, sparse synapses (such as chromaffin cells as reported in Logan et al. [43]) and fast imaging on the order of 50 Hz to capture fleeting recycling events such as kiss-and-run. In addition to the use of mCLING as a vesicle endocytosis marker, these imaging experiments could also be performed using FM dyes, which have been successfully used to image endo and exocytosis events in cells [95]. FM dyes are captured from the plasma membrane during vesicle uptake, and they diffuse into the extracellular space during exocytosis where they can be washed



away with saline. Observing a-syn's and VAMP-2's colocalisation with internalised vesicles labelled with FM dye could validate the results presented in this chapter, confirming it is not an effect specific to mCLING labelling.

### **3.7.2 A-syn and VAMP-2 clusters increase in size in the absence of calcium**

Dopaminergic neurons, the most vulnerable neurons in Parkinson's Disease, undergo calcium fluctuations throughout their lifetime [48, 49]. Calcium is a key regulator of vesicle recycling, as it provides cues for the release or uptake of neurotransmitters [50]. Calcium has been shown to mediate the lipid-binding activity of a-syn's C terminus [17, 18, 5], and to affect a-syn's clustering behaviour in synaptosomes [5]. Lautenschlager et al. [5] argued that in the presence of calcium, a-syn may act as an extended double anchor to tether synaptic vesicles together, causing them to cluster 3.3. In their study, a-syn clusters increased in size upon calcium depletion using EGTA, whereas VAMP-2 clusters showed no difference in size in the presence or absence of calcium. In this study, both a-syn and VAMP-2 distributions were observed to increase in size in the absence of calcium, with a slightly larger increase for VAMP-2 than for a-syn (Fig. 3.26). The analysis presented in this chapter includes more than twice as many synaptosomes as in [5], and corroborates the observed effect using two cluster analysis methods.

Burre et al. [41] have previously shown that a-syn, via its C-terminus, binds to the N-terminus of VAMP-2 on synaptic vesicle surfaces. More recently, Sun et al. [42] proposed an "inter-locking" model in which a-syn binds to VAMP-2 in adjacent vesicles to maintain physiological synaptic vesicle clustering. This interaction between a-syn and VAMP-2 could potentially explain the cluster size increase observed here for both proteins, upon calcium depletion. Alternatively, this could be due to both a-syn and VAMP-2 having a calcium-dependent localisation at the presynapse.

### **3.7.3 SynaptoAnalysis is a useful tool for analysing multi-colour SMLM data from synaptosomes**

SMLM data sets can be very large, occupying a few gigabytes per imaged FOV. When imaging 5-10 different FOVs for multiple experimental conditions (temperature, or calcium), each with two or three repeats, data sets can occupy hundreds of gigabytes. The reconstruction and analysis of these data sets can be time and computationally intensive, and it can be cumbersome to analyse with available colocalisation or cluster analysis software.

Instead, a specialised processing (chromatic correction, filtering, segmentation) and analysis (colocalisation and cluster size) software pipeline was built, SynaptoAnalysis, which allowed us to detect hundreds of synaptosomes from *d*STORM imaging data, and to analyse the colocalisation and clustering behaviour of fluorescent markers using either spatial point pattern or image methods.

### Perspectives and future work

As an alternative to the multi-step filtering and segmentation steps in SynaptoAnalysis, Voronoi tessellation could be applied as described by Levet et al. [96] to compartmentalise and classify localisations. Correlations can be determined between the different polygons in the Voronoi diagram, and then used to perform colocalisation and cluster analysis.

To verify the results provided by the Manders individual coefficients during colocalisation analysis, Ripley's bivariate H-function will be implemented in a future SynaptoAnalysis release to define an average interaction distance between the mCLING/a-syn and mCLING/VAMP-2 clusters.

Regarding cluster analysis, the simulations performed on circular point patterns and inhomogeneous point patterns from experimental synaptosome outlines showed the RMSD is insensitive to cluster density and point jitter from localisation uncertainty. In the benchmarking experiments using ground truth synaptosomes (Fig. 3.22) the RMSD underestimated the true radius by a factor of  $\sim 2$ , whereas Ripley's H-function provided a much more accurate estimate of cluster radius. Both metrics were sensitive enough to detect size differences between mCLING clusters closer to the inner membrane of vesicles and a-syn and VAMP-2 being tethered with long antibody linkers to the outer membranes. The differences detected by the RMSD were closer to the true difference expected from the antibody linkers, than that from Ripley's H-function.

In summary, it is useful and important to corroborate cluster size measurements using multiple methods to highlight the shortcomings and advantages of each. For future work with large data sets comparing multiple experimental conditions, i.e. temperature and calcium levels, a machine learning approach as proposed by Williamson et al. [97] could be implemented to classify single-molecule clusters and extract features with no *a priori* knowledge of cluster size, shape, or density.

### 3.7.4 Synthetic SMLM data from a model synaptosome is useful to benchmark cluster analysis methods

Synthetic SMLM data from synaptosomes were generated using the software testSTORM, which allowed us to input parameters to accurately simulate the experimental image acquisition process, and to model a synaptosome stained for mCLING, a-syn, and VAMP-2 using physical characteristics from literature.

As mentioned in the previous section, the simulated synaptosome data sets highlighted important differences between the RMSD and Ripley's H-function in their accuracy to measure cluster sizes (Fig. 3.22). Ripley's H-function approximated more closely the true value (~500 nm) of the modelled synaptosome radius, however the RMSD was more sensitive to the differences in distribution size of the mCLING to that of a-syn and VAMP-2 due to their antibody linker lengths. Moreover, the cluster radius from Ripley's H-function in the synthetic data (a-syn ~596 nm VAMP-2 ~593 nm) were within the standard deviation of the cluster sizes calculated from the experimental data (a-syn =  $451 \pm 142$  nm, VAMP-2 =  $496 \pm 160$  nm), indicating the size distributions in the model are a valid approximation to our experimental data. Contrary to what was previously reported by Kiskowski et al. [85] as mentioned in section 2.3.2 of Chapter 2, Ripley's H-function did not overestimate the cluster radius in either the experimental or simulated data sets measured in this chapter. The bias reported by Kiskowski et al. in measuring domain size with Ripley's H-function could likely stem from their analysis only on low-density patterns, as mentioned in the discussion section of their seminal paper [85]. The high accuracy of Ripley's H-function in estimating a mean cluster radius for our data sets is likely due to the high localisation density in all clusters analysed, as well as to the large number of clusters analysed which compensates for variations in individual measurements.

For colocalisation analysis, it was more difficult to generate representative models of the experimental data set. The simulation assumes perfect spheres with homogenous background, when synaptosomes are really asymmetric spheroid-like blobs with inhomogeneous background signals. Moreover, the measurement of a 2D projection from a 3D simulated structure introduces potential errors, e.g. two structures which are spatially separated in the z direction but share the same x,y coordinates will appear colocalised. This apparent 2D colocalisation of non-overlapping objects in 3D will artificially inflate the measured 2D colocalisation values. A 3D imaging method would be required to access the z information from the fluorescent emitters and examine their colocalisation in 3D, such as that used in Dani et al. [62]. Alternatively, another way to discriminate between fluorescent emitters in the z-direction would be to illuminate the sample using TIRF, thereby reducing the excitation and imaging volume to a 100 nm region above the focal plane (which is equivalent to the axial

penetration depth of the evanescent wave present in TIRF). Additionally, several parameters in the simulation are approximations, such as the number of mCLING molecules per vesicle, or the percentage of successfully labelled binding sites on the vesicles. These quantities are unknown and were impractical to measure with our available instrumentation.

### **3.7.5 Limitations of multi-colour dSTORM to study protein distributions in synaptosomes**

The resolution required to study single synaptic vesicles and their protein content in synaptosomes is currently beyond that offered by super-resolution fluorescence microscopy techniques, due to the dense clustering of the vesicles, leading to too many overlapping fluorophores within a diffraction-limited region. However, in this study SMLM was useful for imaging clusters of vesicles, and their sub-diffractive distributions.

In dSTORM, the photo switching properties of dyes depend strongly on buffer composition; therefore, buffers are often optimised for specific dyes to achieve the best image quality for a single colour channel [79]. In the multi-colour dSTORM experiments in this chapter, the 488 and 561 nm dyes exhibited less optimal switching than the 647 nm channel, even though the buffer used has been reported to work best across all fluorescent dyes [79]. Sub-optimal photo switching reduces the localisation precision achieved and worsens the image quality of reconstructions in our experiments. To account for these shortcomings, strict filters were implemented in our software SynptoAnalysis post-reconstruction to remove sub-optimal localisations prior to analysis. For some data sets, this resulted in a large rejection rate of the captured localisations as exemplified in Fig. 3.14 a and c, in which ~50% and ~75% of the detected synaptosome candidates were rejected by the spatial overlap filter, respectively. These two rejection rates are representative for synaptosome analysis, as they describe the rejection of potentially useful synaptosome candidates after noise and background contributions had already been removed with previous filters. This rejection rate was considered acceptable in these experiments, as we deemed it more important to remove false positives that could bias our biological interpretation than to keep a larger percentage of our data.

### **Future work**

In future studies, to address the shortcomings of dSTORM for multi-colour imaging of protein distributions in synapses, DNA-PAINT could be implemented to decouple the localisation procedure from the individual dye photo physics [98]. This would ensure that the blinking properties and the localisation precision would be similar for all markers and

would reduce the rejection rate of sub-optimal localisations due to unequal photo switching. Moreover, additional colour channels could be added for markers that could further inform our experiments, e.g. a fluorescent calcium channel marker could be added to ascertain if a-syn clusters near or away from calcium channels at physiological vs calcium depleted levels, or an active zone marker such as bassoon to confirm a-syn's location at the active zone. DNA-PAINT, which uses transiently binding oligonucleotide chains to generate the blinking effect required for SMLM, also offers the possibility of performing stoichiometric counting on fluorescent markers in its qPAINT version [99]. Since the blinking effect is only dependent on the binding kinetic of the oligonucleotide pairs, it is independent of dye photophysics and can therefore be used for quantitative measurements of molecule numbers across multiple wavelengths. Implementing DNA-PAINT for this study would require optimization of the labelling procedure for a-syn, VAMP-2, and mCLING using DNA oligomers as docking strands.

To accurately map the location of a-syn and VAMP-2 at the synapse with respect to internalised vesicles, 3D SMLM could be implemented with *d*STORM or DNA-PAINT. The high density of emitters in a small area would make it difficult to obtain information from large 3D engineered PSFs such as the tetrapod [100] or the double helix [101]; therefore, an astigmatic PSF would likely be the best solution to implement. To reduce the density of emitters per volume, and increase the imaging resolution, expansion microscopy [93] could be combined with SMLM as shown by Zwettler et al. [102] to produce 3D reconstructions of a-syn and VAMP-2 in synaptosomes with single vesicle level resolution.

### 3.8 Acknowledgment of contributions

This project is a collaborative and ongoing effort between the Laser Analytics Group and the Molecular Neuroscience Group. The individual contributions for each part of the project described in this chapter are mentioned below. Dr. Ana Fernandez-Villegas performed the rat dissections, and Dr. Amberley Stephens performed the synaptosome preparations and the fluorescent staining, as well as the plating of the samples on multi-well glass dishes for imaging. Tibor Novak from the University of Szeged modified the testSTORM code to include the synaptosome filled with vesicles as a model to perform the simulations. Ezra Bruggeman from KU Leuven and myself wrote SynaptoAnalysis, using parts of previous code written by Dr. Romain Laine and Dr. Florian Ströhl as guides. I performed the *d*STORM imaging and subsequent analysis of the imaging data collected, as well as all characterisation experiments and simulations described. Dr. Gabriele Kaminski Schierle and Dr. Amberley Stephens devised the biological question. Dr. Miklos Erdélyi provided useful guidance in the

development of the simple computer model for the synaptosomes. Dr. Steven Lee, Dr. Eric J Rees, and Professor Clemens Kaminski provided critical advice on the super-resolution imaging procedure and the analysis of the single-molecule data.

## References

- [1] AbdulRasheed A. Alabi and Richard W. Tsien. Perspectives on Kiss-and-Run: Role in Exocytosis, Endocytosis, and Neurotransmission. *Annual Review of Physiology*, 75(1):393–422, 2011.
- [2] Sreeganga Chandra, Francesco Fornai, Hyung-Bae Kwon, Umar Yazdani, Deniz Atasoy, Xinran Liu, Robert E. Hammer, Giuseppe Battaglia, Dwight C. German, Pablo E. Castillo, and Thomas C. Sudhof. Double-knockout mice for alpha- and beta-synucleins: Effect on synaptic functions. *Proceedings of the National Academy of Sciences*, 101(41):14966–14971, 2004.
- [3] Maria Grazia Spillantini, Marie Luise Schmidt, Virginia M.-Y. Lee, John Q Trojanowski, Ross Jakes, and Michel Goedert.  $\alpha$ -Synuclein in Lewy bodies. *Nature*, 388:839–840, 1997.
- [4] Céline Galvagnion, James W. P. Brown, Myriam M. Oubrai, Patrick Flagmeier, Michele Vendruscolo, Alexander K. Buell, Emma Sparr, and Christopher M. Dobson. Chemical properties of lipids strongly affect the kinetics of the membrane-induced aggregation of  $\alpha$ -synuclein. *Proceedings of the National Academy of Sciences*, 113(26):7065–7070, 2016.
- [5] Janin Lautenschläger, Amberley D. Stephens, Giuliana Fusco, Florian Ströhl, Nathan Curry, Maria Zacharopoulou, Claire H. Michel, Romain Laine, Nadezhda Nespovitaya, Marcus Fantham, Dorothea Pinotsi, Wagner Zago, Paul Fraser, Anurag Tandon, Peter St George-Hyslop, Eric Rees, Jonathan J. Phillips, Alfonso De Simone, Clemens F. Kaminski, and Gabriele S. Kaminski Schierle. C-terminal calcium binding of  $\alpha$ -synuclein modulates synaptic vesicle interaction. *Nature Communications*, 9(1):712, 2018.
- [6] Janin Lautenschläger, Clemens F. Kaminski, and Gabriele S. Kaminski Schierle.  $\alpha$ -Synuclein – Regulator of Exocytosis, Endocytosis, or Both? *Trends in Cell Biology*, 27(7):468–479, 2017.

- [7] Christina R. Bodner, Christopher M. Dobson, and Ad Bax. Multiple Tight Phospholipid-Binding Modes of  $\alpha$ -Synuclein Revealed by Solution NMR Spectroscopy. *Journal of Molecular Biology*, 390(4):775–790, 2009.
- [8] Sreeganga Chandra, Xiaocheng Chen, Josep Rizo, Reinhard Jahn, and Thomas C. Südhof. A broken  $\alpha$ -helix in folded  $\alpha$ -synuclein. *Journal of Biological Chemistry*, 278(17):15313–15318, 2003.
- [9] Elka R. Georgieva, Trudy F. Ramlall, Peter P. Borbat, Jack H. Freed, and David Eliezer. Membrane-bound  $\alpha$ -synuclein forms an extended helix: Long-distance pulsed ESR measurements using vesicles, bicelles, and rodlike micelles. *Journal of the American Chemical Society*, 130(39):12856–12857, 2008.
- [10] C. C. Jao, B. G. Hegde, J. Chen, I. S. Haworth, and R. Langen. Structure of membrane-bound alpha-synuclein from site-directed spin labeling and computational refinement. *Proceedings of the National Academy of Sciences*, 105(50):19666–19671, 2008.
- [11] Elizabeth R. Middleton and Elizabeth Rhoades. Effects of curvature and composition on  $\alpha$ -synuclein binding to lipid vesicles. *Biophysical Journal*, 99(7):2279–2288, 2010.
- [12] Adam J. Trexler and Elizabeth Rhoades. Alpha-synuclein binds large unilamellar vesicles as an extended helix- Supplementary Info. *Biochemistry*, 48(11):1–6, 2009.
- [13] Tobias S. Ulmer, Ad Bax, Nelson B. Cole, and Robert L. Nussbaum. Structure and dynamics of micelle-bound human  $\alpha$ -synuclein. *Journal of Biological Chemistry*, 280(10):9595–9603, 2005.
- [14] W S Davidson, A Jonas, D F Clayton, and J M George. Stabilization of alpha-synuclein secondary structure upon binding to synthetic membranes. *The Journal of biological chemistry*, 273(16):9443–9, apr 1998.
- [15] Jobin Varkey, Jose Mario Isas, Naoko Mizuno, Martin Borch Jensen, Vikram Kjøller Bhatia, Christine C. Jao, Jitka Petrova, John C. Voss, Dimitrios G. Stamou, Alasdair C. Steven, and Ralf Langen. Membrane curvature induction and tubulation are common features of synucleins and apolipoproteins. *Journal of Biological Chemistry*, 285(42):32486–32493, 2010.
- [16] Christopher H. Westphal and Sreeganga S. Chandra. Monomeric synucleins generate membrane curvature. *Journal of Biological Chemistry*, 288(3):1829–1840, 2013.

- [17] Giuliana Fusco, Alfonso De Simone, Paolo Arosio, Michele Vendruscolo, Gianluigi Veglia, and Christopher M. Dobson. Structural Ensembles of Membrane-bound  $\alpha$ -Synuclein Reveal the Molecular Determinants of Synaptic Vesicle Affinity. *Scientific Reports*, 6(May):1–9, 2016.
- [18] Giuliana Fusco, Alfonso De Simone, Tata Gopinath, Vitaly Vostrikov, Michele Vendruscolo, Christopher M. Dobson, and Gianluigi Veglia. Direct observation of the three regions in  $\alpha$ -synuclein that determine its membrane-bound behaviour. *Nature Communications*, 5(May):1–8, 2014.
- [19] D D Murphy, S M Rueter, J Q Trojanowski, and V M Lee. Synucleins are developmentally expressed, and alpha-synuclein regulates the size of the presynaptic vesicular pool in primary hippocampal neurons. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 20(9):3214–3220, 2000.
- [20] Karina J Vargas, Sachin Makani, Taylor Davis, Christopher H Westphal, Pablo E Castillo, and Sreeganga S Chandra. Synucleins Regulate the Kinetics of Synaptic Vesicle Endocytosis. 34(28):9364–9376, 2014.
- [21] Seung J. Lee, Hyesung Jeon, and Konstantin V. Kandrор.  $\alpha$ -Synuclein is localized in a subpopulation of rat brain synaptic vesicles. *Acta Neurobiologiae Experimentalis*, 68(4):509–515, 2008.
- [22] Quan Gan and Shigeki Watanabe. Synaptic vesicle endocytosis in different model systems. *Frontiers in Cellular Neuroscience*, 12:171, jun 2018.
- [23] Yasunori Saheki and Pietro De Camilli. Synaptic vesicle endocytosis. *Cold Spring Harbor Perspectives in Biology*, 7(3):323–330, 2012.
- [24] Marko Kaksonen and Aurélien Roux. Mechanisms of clathrin-mediated endocytosis. *Nature Reviews Molecular Cell Biology*, 19(5):313–326, 2018.
- [25] Natali L. Chanaday and Ege T. Kavalali. Time course and temperature dependence of synaptic vesicle endocytosis. *FEBS Letters*, 592(21):3606–3614, 2018.
- [26] Axel T. Brunger, Ucheor B. Choi, Ying Lai, Jeremy Leitz, and Qiangjun Zhou. Molecular Mechanisms of Fast Neurotransmitter Release. *Annual Review of Biophysics*, 47:469–97, 2018.
- [27] Satyan Sharma and Manfred Lindau. The mystery of the fusion pore. *Nature Structural and Molecular Biology*, 23(1):5–6, 2016.



- [28] K. E. Larsen, Y. Schmitz, M. D. Troyer, E. Mosharov, P. Dietrich, A. Z. Quazi, M. Savalle, V. Nemani, F. A. Chaudhry, R. H. Edwards, L. Stefanis, and D. Sulzer.  $\alpha$ -Synuclein Overexpression in PC12 and Chromaffin Cells Impairs Catecholamine Release by Interfering with a Late Step in Exocytosis. *Journal of Neuroscience*, 26(46):11915–11922, 2006.
- [29] Venu M. Nemani, Wei Lu, Victoria Berge, Ken Nakamura, Bibiana Onoa, Michael K. Lee, Farrukh A. Chaudhry, Roger A. Nicoll, and Robert H. Edwards. Increased Expression of  $\alpha$ -Synuclein Reduces Neurotransmitter Release by Inhibiting Synaptic Vesicle Reclustering after Endocytosis. *Neuron*, 65(1):66–79, 2010.
- [30] D. A. Scott, I. Tabarean, Y. Tang, A. Cartier, E. Masliah, and S. Roy. A Pathologic Cascade Leading to Synaptic Dysfunction in  $\alpha$ -Synuclein-Induced Neurodegeneration. *Journal of Neuroscience*, 30(24):8083–8095, 2010.
- [31] Irina Gureviciene, Kestutis Gurevicius, and Heikki Tanila. Role of  $\alpha$ -synuclein in synaptic glutamate release. *Neurobiology of Disease*, 28(1):83–89, 2007.
- [32] Stephanie Janezic, Sarah Threlfell, Paul D Dodson, Megan J Dowie, Tonya N Taylor, Dawid Potgieter, Laura Parkkinen, Steven L Senior, Sabina Anwar, Brent Ryan, Thierry Deltheil, Polina Kosillo, Milena Cioroch, Katharina Wagner, Olaf Ansorge, David M Bannerman, J Paul Bolam, Peter J Magill, Stephanie J Cragg, and Richard Wade-Martins. Deficits in dopaminergic transmission precede neuron loss and dysfunction in a new Parkinson model. *Proceedings of the National Academy of Sciences of the United States of America*, 110(42):E4016–25, 2013.
- [33] Nanping Wu, Prasad R. Joshi, Carlos Cepeda, Eliezer Masliah, and Michael S. Levine. Alpha-Synuclein overexpression in mice alters synaptic communication in the corticostriatal pathway. *Journal of Neuroscience Research*, 88(8):1764–1776, 2010.
- [34] Shumin Liu, Ipe Ninan, Irina Antonova, Fortunato Battaglia, Fabrizio Trinchese, Archana Narasanna, Nikolai Kolodilov, William Dauer, Robert D. Hawkins, and Ottavio Arancio.  $\alpha$ -Synuclein Produces a Long-Lasting Increase in Neurotransmitter Release. *EMBO Journal*, 23(22):4506–4516, 2004.
- [35] Deborah E Cabin, Kazuhiro Shimazu, Diane Murphy, Nelson B Cole, Wolfram Gottschalk, Kellie L McIlwain, Bonnie Orrison, Amy Chen, Christopher E Ellis, Richard Paylor, Bai Lu, and Robert L Nussbaum. Synaptic vesicle depletion correlates with attenuated synaptic responses to prolonged repetitive stimulation in mice lacking

- alpha-synuclein. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 22(20):8797–807, 2002.
- [36] Eduardo D. Martín, Carmen González-García, Miguel Milán, Isabel Fariñas, and Valentín Ceña. Stressor-related impairment of synaptic transmission in hippocampal slices from alpha-synuclein knockout mice. *European Journal of Neuroscience*, 20(11):3085–3091, 2004.
- [37] J. B. Watson, A. Hatami, H. David, E. Masliah, K. Roberts, C. E. Evans, and M. S. Levine. Alterations in corticostriatal synaptic plasticity in mice overexpressing human  $\alpha$ -synuclein. *Neuroscience*, 159(2):501–513, 2009.
- [38] Asa Abeliovich, Yvonne Schmitz, Isabel Fariñ, Derek Choi-Lundberg, Wei-Hsien Ho, Pablo E Castillo, Natasha Shinsky, Jose Manuel, Garcia Verdugo, Mark Armanini, Anne Ryan, Mary Hynes, and Heidi Phillips. Mice Lacking Alpha-Synuclein Display Functional Deficits in the Nigrostriatal Dopamine System. *Neuron*, 25:239–252, 2000.
- [39] Steven L. Senior, Natalia Ninkina, Robert Deacon, David Bannerman, Vladimir L. Buchman, Stephanie J. Cragg, and Richard Wade-Martins. Increased striatal dopamine release and hyperdopaminergic-like behaviour in mice lacking both alpha-synuclein and gamma-synuclein. *European Journal of Neuroscience*, 27(4):947–957, 2008.
- [40] Leonid Yavich, Heikki Tanila, Saila Vepsäläinen, and Pekka Jäkälä. Role of alpha-Synuclein in Presynaptic Dopamine Recruitment. *Journal of Neuroscience*, 24(49):11165–11170, 2004.
- [41] Jacqueline Burré, Manu Sharma, Theodoros Tsetsenis, Vladimir Buchman, Mark R Etherton, and Thomas C Südhof. Alpha-Synuclein Promotes SNARE-Complex Assembly in Vivo and in Vitro. *Science*, 1663(2010):1663–1668, 2010.
- [42] Jichao Sun, Lina Wang, Huan Bao, Sanjay Premi, Utpal Das, Edwin R. Chapman, and Subhojit Roy. Functional cooperation of  $\alpha$ -synuclein and VAMP2 in synaptic vesicle recycling. *Proceedings of the National Academy of Sciences*, 116(23):201903049, 2019.
- [43] Todd Logan, Jacob Bendor, Chantal Toupin, Kurt Thorn, and Robert H. Edwards.  $\alpha$ -Synuclein promotes dilation of the exocytotic fusion pore. *Nature Neuroscience*, 20(5):681–689, 2017.

- [44] Tziona Ben Gedalya, Virginie Loeb, Eitan Israeli, Yoram Altschuler, Dennis J. Selkoe, and Ronit Sharon.  $\alpha$ -Synuclein and polyunsaturated fatty acids promote clathrin-mediated endocytosis and synaptic vesicle recycling. *Traffic*, 10(2):218–234, 2009.
- [45] Furong Cheng, Xin Li, Yaohua Li, Chaodong Wang, Tao Wang, Guangwei Liu, Andrius Baskys, Kenji Ueda, Piu Chan, and Shun Yu.  $\alpha$ -Synuclein promotes clathrin-mediated NMDA receptor endocytosis and attenuates NMDA-induced dopaminergic cell death. *Journal of Neurochemistry*, 119(4):815–825, 2011.
- [46] Haya Kisos, Tziona Ben-Gedalya, and Ronit Sharon. The clathrin-dependent localization of dopamine transporter to surface membranes is affected by  $\alpha$ -Synuclein. *Journal of Molecular Neuroscience*, 52(2):167–176, 2014.
- [47] Shawn M Ferguson and Pietro De Camilli. Dynamin, a membrane-remodelling GTPase. *Nature reviews. Molecular cell biology*, 13(2):75–88, 2012.
- [48] Michael J. Hurley and David T. Dexter. Voltage-gated calcium channels and Parkinson's disease. *Pharmacology and Therapeutics*, 133(3):324–333, 2012.
- [49] Weifeng Xu and Diane Lipscombe. Neuronal Cav1.3 $\alpha$ 1 L-type channels activate at relatively hyperpolarized membrane potentials and are incompletely inhibited by dihydropyridines. *Journal of Neuroscience*, 21(16):5944–5951, 2001.
- [50] Jerome Di Giovanni, Cé Cile Iborra, Yves Maulet, Christian Lévêque, Oussama El Far, and Michael Seagar. Calcium-dependent regulation of SNARE-mediated membrane fusion by calmodulin. *Journal of Biological Chemistry*, 285(31):23665–23675, 2010.
- [51] Natalia H. Revelo, Dirk Kamin, Sven Truckenbrodt, Aaron B. Wong, Kirsten Reuter-Jessen, Ellen Reisinger, Tobias Moser, and Silvio O. Rizzoli. A new probe for super-resolution imaging of membranes elucidates trafficking pathways. *Journal of Cell Biology*, 205(4):591–606, 2014.
- [52] Christopher J Rowlands, Florian Ströhl, Pedro P Vallejo Ramirez, Katharina M Scherer, and Clemens F Kaminski. Flat-Field Super-Resolution Localization Microscopy with a Low-Cost Refractive Beam-Shaping Element. *Scientific Reports*, 8(1):5630, 2018.
- [53] Kyle M. Douglass, Christian Sieben, Anna Archetti, Ambroise Lambert, and Sulliana Manley. Super-resolution imaging of multiple cells by optimized flat-field epi-illumination. *Nature Photonics*, 10(11):705–708, 2016.

- [54] Makio Tokunaga, Naoko Imamoto, and Kumiko Sakata-Sogawa. Highly inclined thin illumination enables clear single-molecule imaging in cells. *Nature methods*, 5(2), 2008.
- [55] Sebastian van de Linde, Anna Löschberger, Teresa Klein, Meike Heidebreder, Steve Wolter, Mike Heilemann, and Markus Sauer. Direct stochastic optical reconstruction microscopy with standard fluorescent probes. *Nature protocols*, 6(7):991–1009, 2011.
- [56] Martin Ovesný, Pavel Křížek, Josef Borkovec, Zdeněk Švindrych, and Guy M. Hagen. ThunderSTORM: A comprehensive ImageJ plug-in for PALM and STORM data analysis and super-resolution imaging. *Bioinformatics*, 30(16):2389–2390, 2014.
- [57] Technoloy Andor. iXon Ultra X-9847 Performance Sheet, Issue 1, 2014.
- [58] Joses Ho, Tayfun Tumkaya, Sameer Aryal, Hyungwon Choi, and Adam Claridge-Chang. Moving beyond P values: data analysis with estimation graphics. *Nature Methods*, 16(7):565–566, 2019.
- [59] Dana Bar-On, Steve Wolter, Sebastian Van De Linde, Mike Heilemann, German Nudelman, Esther Nachliel, Menachem Gutman, Markus Sauer, and Uri Ashery. Super-resolution imaging reveals the internal architecture of nano-sized syntaxin clusters. *Journal of Biological Chemistry*, 287(32):27158–27167, 2012.
- [60] Carolin Böger, Anne-Sophie Hafner, Thomas Schlichthärle, Maximilian T. Strauss, Sebastian Malkusch, Ulrike Endesfelder, Ralf Jungmann, Erin M. Schuman, and Mike Heilemann. Super-resolution imaging and estimation of protein copy numbers at single synapses with DNA-point accumulation for imaging in nanoscale topography. *Neurophotonics*, 6(03):1, 2019.
- [61] Ingrid Chamma, Florian Levet, Jean-Baptiste Sibarita, Matthieu Sainlos, and Olivier Thoumine. Nanoscale organization of synaptic adhesion proteins revealed by single-molecule localization microscopy. *Neurophotonics*, 3(4):041810, 2016.
- [62] Adish Dani, Bo Huang, Joseph Bergan, Catherine Dulac, and Xiaowei Zhuang. Superresolution Imaging of Chemical Synapses in the Brain. *Neuron*, 68(5):843–856, 2010.
- [63] Nadine Ehmann, Sebastian Van De Linde, Amit Alon, Dmitrij Ljaschenko, Xi Zhen Keung, Thorge Holm, Annika Rings, Aaron DiAntonio, Stefan Hallermann, Uri Ashery, Manfred Heckmann, Markus Sauer, and Robert J. Kittel. Quantitative super-resolution

- imaging of Bruchpilot distinguishes active zone states. *Nature Communications*, 5, 2014.
- [64] Syuan Ming Guo, Remi Veneziano, Simon Gordonov, Li Li, Eric Danielson, Karen Perez de Arce, Demian Park, Anthony B. Kulesa, Eike Christian Wamhoff, Paul C. Blainey, Edward S. Boyden, Jeffrey R. Cottrell, and Mark Bathe. Multiplexed and high-throughput neuronal fluorescence imaging with diffusible probes. *Nature Communications*, 10(1), 2019.
- [65] Kalina T. Haas, Benjamin Compans, Mathieu Letellier, Thomas M. Bartol, Dolors Grillo-Bosch, Terrence J. Sejnowski, Matthieu Sainlos, Daniel Choquet, Olivier Thoumine, and Eric Hosy. Pre-post synaptic alignment through neuroligin-1 tunes synaptic transmission efficiency. *eLife*, 7:1–22, 2018.
- [66] Janosch Peter Heller, Tuamoru Odii, Kaiyu Zheng, and Dmitri A. Rusakov. Imaging tripartite synapses using super-resolution microscopy. *Methods*, 174(May 2019):81–90, 2020.
- [67] Harold D. MacGillavry, Yu Song, Sridhar Raghavachari, and Thomas A. Blanpied. Nanoscale scaffolding domains within the postsynaptic density concentrate synaptic ampa receptors. *Neuron*, 78(4):615–622, 2013.
- [68] Deepak Nair, Eric Hosy, Jennifer D. Petersen, Audrey Constals, Gregory Giannone, Daniel Choquet, and Jean Baptiste Sibarita. Super-resolution imaging reveals that AMPA receptors inside synapses are dynamically organized in nanodomains regulated by PSD95. *Journal of Neuroscience*, 33(32):13204–13224, 2013.
- [69] Francesca Pennacchietti, Sebastiano Vascon, Thierry Nieuu, Christian Rosillo, Sabyasachi Das, Shiva K. Tyagarajan, Alberto Diaspro, Alessio del Bue, Enrica Maria Petrini, Andrea Barberis, and Francesca Cella Zanacchi. Nanoscale molecular reorganization of the inhibitory postsynaptic density is a determinant of gabaergic synaptic potentiation. *Journal of Neuroscience*, 37(7):1747–1756, 2017.
- [70] Hirokazu Sakamoto, Tetsuroh Ariyoshi, Naoya Kimpara, Kohtaroh Sugao, Isamu Taiko, Kenji Takikawa, Daisuke Asanuma, Shigeyuki Namiki, and Kenzo Hirose. Synaptic weight set by Munc13-1 supramolecular assemblies. *Nature Neuroscience*, 21(1):41–55, 2018.
- [71] Christian G. Specht, Ignacio Izeddin, Pamela C. Rodriguez, Mohamed ElBeheiry, Philippe Rostaing, Xavier Darzacq, Maxime Dahan, and Antoine Triller. Quantitative

- nanoscopy of inhibitory synapses: Counting gephyrin molecules and receptor binding sites. *Neuron*, 79(2):308–321, 2013.
- [72] Ai Hui Tang, Haiwen Chen, Tuo P. Li, Sarah R. Metzbower, Harold D. MacGillavry, and Thomas A. Blanpied. A trans-synaptic nanocolumn aligns neurotransmitter release to receptors. *Nature*, 536(7615):210–214, 2016.
- [73] Gareth J.O. Evans. The synaptosome as a model system for studying synaptic physiology. *Cold Spring Harbor Protocols*, 2015(5):421–424, 2015.
- [74] Nobuyuki Otsu. A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, C(1):62–66, 1979.
- [75] Paolo Annibale, Marco Scarselli, Mattia Greco, and Aleksandra Radenovic. Identification of the factors affecting co-localization precision for quantitative multicolor localization microscopy. *Optical Nanoscopy*, 1(1):1–13, 2012.
- [76] Ardeshir Goshtasby. Image registration by local approximation methods. *Image vision and Computing*, 6(4), 1988.
- [77] L Stirling Churchman, Ökten Zeynep, Ronald S Rock, John F Dawson, and James A Spudich. Single molecule high-resolution colocalization of Cy3 and Cy5 attached to macromolecules measures intramolecular distances through time. *Proceedings of the National Academy of Sciences*, 102(5):1419–1423, 2005.
- [78] Peter R. Dunkley, Paula E. Jarvie, and Phillip J. Robinson. A rapid percoll gradient procedure for preparation of synaptosomes. *Nature Protocols*, 3(11):1718–1728, 2008.
- [79] Graham T. Dempsey, Joshua C. Vaughan, Kok Hao Chen, Mark Bates, and Xiaowei Zhuang. Evaluation of fluorophores for optimal performance in localization-based super-resolution imaging. *Nature Methods*, 8(12):1027–1040, 2011.
- [80] Benjamin G. Wilhelm, Sunit Mandad, Sven Truckenbrodt, Katharina Kröhnert, Christina Schäfer, Burkhard Rammner, Seong Joo Koo, Gala A Claßen, Michael Krauss, Volker Haucke, Henning Urlaub, and Silvio O Rizzoli. Vesicle Trafficking Proteins. *Science*, 344(6187):1023–1028, 2014.
- [81] E.M.M; Manders, F. J. Verbeek, and J. A. Aten. Measurement of co-localization of objects in dual-colour confocal images, 1993.

- [82] Fabrice P. Cordelières and Susanne Bolte. Experimenters' guide to colocalization studies. Finding a way through indicators and quantifiers, in practice. *Methods in Cell Biology*, 123:395–408, 2014.
- [83] Kenneth W Dunn, Malgorzata M Kamocka, and John H McDonald. A practical guide to evaluating colocalization in biological microscopy. *American Journal of Physiology-Cell Physiology*, 300(4):723–742, 2011.
- [84] Thibault Lagache, Alexandre Grassart, Stéphane Dallongeville, Orestis Faklaris, Nathalie Sauvonnet, Alexandre Dufour, Lydia Danglot, and Jean Christophe Olivo-Marin. Mapping molecular assemblies with fluorescence microscopy and object-based spatial statistics. *Nature Communications*, 9(1):102–108, 2018.
- [85] Maria A. Kiskowski, John F. Hancock, and Anne K. Kenworthy. On the use of Ripley's K-function and its derivatives to analyze domain size. *Biophysical Journal*, 97(4):1095–1103, 2009.
- [86] Shigeo Takamori, Matthew Holt, Katinka Stenius, Edward A. Lemke, Mads Grønborg, Dietmar Riedel, Henning Urlaub, Stephan Schenck, Britta Brügger, Philippe Ringler, Shirley A. Müller, Burkhard Rammner, Frauke Gräter, Jochen S. Hub, Bert L. De Groot, Gottfried Mieskes, Yoshinori Moriyama, Jürgen Klingauf, Helmut Grubmüller, John Heuser, Felix Wieland, and Reinhard Jahn. Molecular Anatomy of a Trafficking Organelle. *Cell*, 127(4):831–846, 2006.
- [87] Tibor Novák, Tamás Gajdos, József Sinkó, Gábor Szabó, and Miklós Erdélyi. Test-STORM: Versatile simulator software for multimodal super-resolution localization fluorescence microscopy. *Scientific Reports*, 7(1):1–8, 2017.
- [88] Amanda M. Hussey and James J. Chambers. Methods to locate and track ion channels and receptors expressed in live neurons. *ACS Chemical Neuroscience*, 6(1):189–198, 2015.
- [89] Mohammad A.A. Fakhree, Niels Zijlstra, Christian C. Raiss, Carolus J. Siero, Heinrich Grabmayr, Andreas R. Bausch, Christian Blum, and Mireille M.A.E. Claessens. The number of  $\alpha$ -synuclein proteins per vesicle gives insights into its physiological function. *Scientific Reports*, 6(1):1–9, 2016.
- [90] Dylan M Owen, Carles Rentero, Astrid Magenau, David Williamson, Macarena Rodriguez, and Katharina Gaus. PALM imaging and cluster analysis of protein heterogeneity at the cell surface. 454(7):446–454, 2010.

- [91] Yu A Chen, Richard H Scheller, and Howard Hughes Medical. SNARE-mediated membrane fusion. *Nature Reviews Molecular Cell Biology*, 2(February):98–106, 2001.
- [92] Wonchul Shin, Lihao Ge, Gianvito Arpino, Seth A. Villarreal, Edaeni Hamid, Huisheng Liu, Wei Dong Zhao, Peter J. Wen, Hsueh Cheng Chiang, and Ling Gang Wu. Visualization of Membrane Pore in Live Cells Reveals a Dynamic-Pore Theory Governing Fusion and Endocytosis. *Cell*, 173(4):934–945.e12, 2018.
- [93] Fei Chen, Paul W Tillberg, and Edward S Boyden. Expansion microscopy. *Science*, 347(6221):543 – 548, 2015.
- [94] Bo-Jui Chang, Victor Didier Perez Meza, and Ernst H. K. Stelzer. csiLSFM combines light-sheet fluorescence microscopy and coherent structured illumination for a lateral resolution below 100 nm. *Proceedings of the National Academy of Sciences*, 2017.
- [95] Michael A Gaffield and William J Betz. Imaging synaptic vesicle exocytosis and endocytosis with FM dyes. *Nature Protocols*, 1(6):2916–2921, 2006.
- [96] Florian Levet, Guillaume Julien, Rémi Galland, Corey Butler, Anne Beghin, Anaël Chazeau, Philipp Hoess, Jonas Ries, Grégory Giannone, and Jean Baptiste Sibarita. A tessellation-based colocalization analysis approach for single-molecule localization microscopy. *Nature Communications*, 10(1):1–12, 2019.
- [97] David J. Williamson, Garth L. Burn, Sabrina Simoncelli, Juliette Griffié, Ruby Peters, Daniel M. Davis, and Dylan M. Owen. Machine learning for cluster analysis of localization microscopy data. *Nature Communications*, 11(1):1–10, 2020.
- [98] Joerg Schnitzbauer, Maximilian T. Strauss, Thomas Schlichthaerle, Florian Schueder, and Ralf Jungmann. Super-resolution microscopy with DNA-PAINT. *Nature Protocols*, 12(6):1198–1228, 2017.
- [99] Ralf Jungmann, Maier S. Avendaño, Mingjie Dai, Johannes B. Woehrstein, Sarit S. Agasti, Zachary Feiger, Avital Rodal, and Peng Yin. Quantitative super-resolution imaging with qPAINT. *Nature Methods*, 13(5):439–442, 2016.
- [100] Yoav Shechtman, Lucien E Weiss, Adam S Backer, Steffen J Sahl, and W E Moerner. Precise Three-Dimensional Scan-Free Multiple-Particle Tracking over Large Axial Ranges with Tetrapod Point Spread Functions. *Nano Letters*, 15(6):4194–4199, 2015.



- 
- [101] Sri Rama Prasanna Pavani, Michael A Thompson, Julie S Biteen, Samuel J Lord, Na Liu, Robert J Twieg, Rafael Piestun, and W E Moerner. Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function. *Proceedings of the National Academy of Sciences*, 106(9):2995–2999, 2009.
- [102] Fabian U Zwettler, Sebastian Reinhard, Davide Gambarotto, Toby D M Bell, Virginie Hamel, Paul Guichard, and Markus Sauer. Molecular resolution imaging by post-labeling expansion single-molecule localization microscopy (Ex-SMLM). *bioRxiv*, 2020.



## **Chapter 4**

# **A study of the colocalisation between HIV and the ESCRT-II protein EAP45**

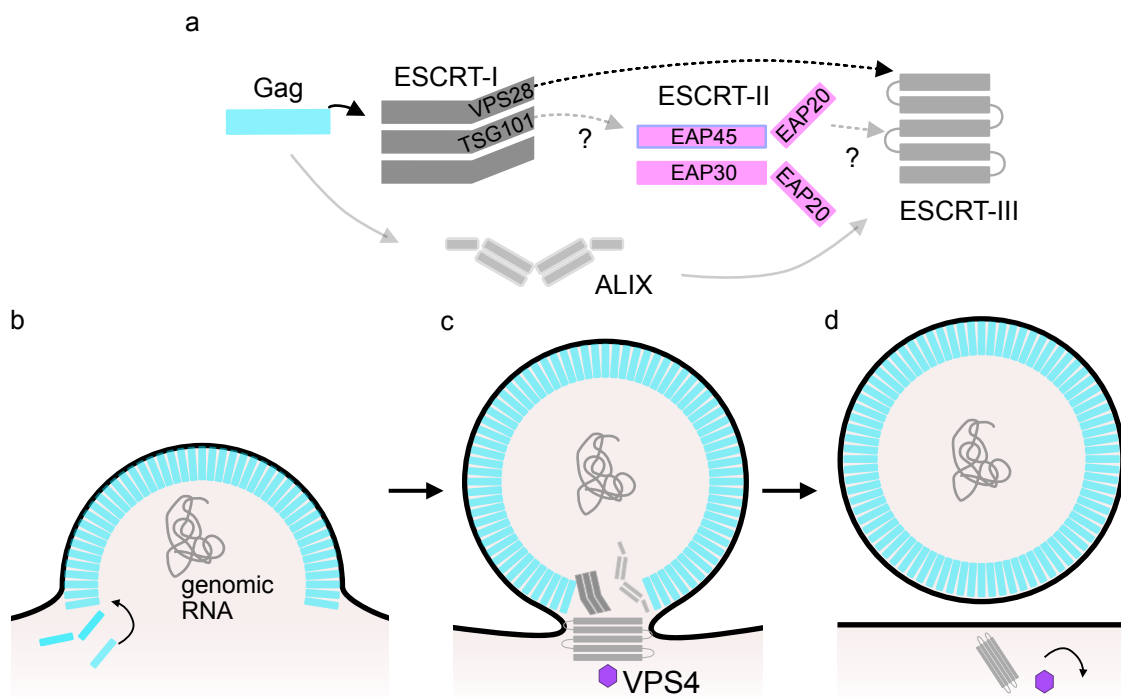
This chapter describes the study the interaction between the endosomal sorting complex required for transport (ESCRT)-II, and the human immunodeficiency virus (HIV) during viral egress using TIRF microscopy and novel imaging analysis tools. The scission of lipid membranes is a common biological process, often mediated by a family of proteins known as the ESCRTs assembling around the separation point. The functions of ESCRT-I and ESCRT-III are well established in certain of these cellular processes, e.g. HIV budding, cytokinesis, and late endosome recruitment; however, the function of ESCRT-II remains unclear. In this study, the main component of ESCRT-II, the protein EAP45, was labelled with a SNAP-tag and imaged using TIRF microscopy to measure its colocalisation with the HIV structural polyprotein Gag. Gag was found to colocalise with EAP45 comparably to ALIX, a protein well-known to be involved in HIV budding, but this colocalisation was significantly reduced upon deletion of the EAP45 N-terminus. These findings strengthen previous evidence showing the importance of the N-terminus of EAP45 for linking to ESCRT-I during HIV budding. In addition, single particle tracking was applied to visualise the dynamic association between EAP45 and Gag, and a method to visualise the concurrent motion of two particles revealed three distinct classes of EAP45 movement relative to Gag. EAP45 particles were found to associate with Gag particles for ~2 mins on average. A preprint of this work is available on bioRxiv [1], and it has been submitted for consideration to a peer-reviewed journal.

## 4.1 Background to HIV and the ESCRT machinery

### 4.1.1 ESCRT proteins are used for membrane remodelling

ESCRTs (endosomal sorting complexes required for transport) are protein complexes involved in a variety of cellular functions, from the scission of the intercellular bridge during cell division (cytokinesis), to the formation of multi-vesicular bodies (MVB) and the budding of HIV. ESCRT activities involve the topological remodelling of membranous structures. The core of the ESCRT machinery is composed of three sub-complexes, ESCRT-I, -II, and -III, which act in a sequentially coordinated manner during membrane scission [2].

Many viruses take advantage of the ESCRT proteins during their life cycle, primarily by hijacking the ESCRT machinery to exit the cell via budding at the plasma membrane [3]. In the case of HIV, Gag polyproteins form clusters in the cytosol before they travel to the plasma membrane, where further Gag molecules accumulate and a bud begins to form [4, 5, 6] (Fig. 4.1 b-d).



**Fig. 4.1** Sequential recruitment of the ESCRTs and ALIX by the polyprotein Gag during HIV budding. (a) The canonical view of ESCRT recruitment involves either recruitment of ESCRT-III via ESCRT-I or ALIX, however the role of ESCRT-II is unclear. (b-d) HIV budding process from the initial accumulation of Gag at the plasma membrane (b), to the recruitment of the ESCRTs at the bud neck to perform the membrane scission (c), followed by the successful budding of the virion from the membrane (d).

### 4.1.2 HIV sequentially recruits ESCRTs during budding

Two distinct motifs of Gag which are active late in the viral budding process, PTAP and YPXL, interact with the early ESCRT-I protein TSG101 and the ESCRT-associated protein ALIX, respectively. These then recruit ESCRT-III for the final scission to occur [6]. This canonical view of the sequential recruitment of the ESCRTs and ALIX by Gag during budding is illustrated in Fig. 4.1 a. Extensive work has been carried out to understand the function of ESCRT-I and ESCRT-III in HIV budding and in other cellular processes [7, 8], however the role of ESCRT-II remains unclear.

ESCRT-II comprises one copy of EAP30 and EAP45 and two copies of EAP20 forming a Y shaped structure [9] (Fig. 4.1 a). Structural studies have shown ESCRT-II interacts with ESCRT-I via the N-terminus of the EAP45 Glue and H0 linker domains, and the C-terminus of the ESCRT-I protein VPS28 [10, 9, 2]. A similar binding region in EAP45 was also shown to interact with the ESCRT-I protein TSG101 [11]. The remaining domains of EAP45, and EAP30, form the core structure of the ESCRT-II complex with EAP20 [12, 9], which is responsible for recruiting ESCRT-III [9].

### 4.1.3 The role of ESCRT-II in HIV budding is unclear

Although its role in linking to ESCRT-I in MVB formation [13] and cytokinesis [14, 15] is accepted, the same linkage has been controversial in HIV budding. An earlier siRNA knockdown (KD) study of EAP20 suggested ESCRT-II is dispensable for HIV budding in 293T cells [11]. However, researchers from the Lever Group (Department of Medicine, Cambridge University) using HAP1 and T cells in which EAP45 had been knocked out (KO) using CRISPR-Cas9 [16] showed that EAP45 is required for efficient HIV budding and spreading [17, 18]. The specificity of this effect was demonstrated by the ability of EAP45 *in trans* to rescue viral budding in EAP45 KO cells, and by its dependency on the H0 linker region in linking to ESCRT-I. The discrepancy between the siRNA KD and the CRISPR/EAP45 KO studies may have arisen from a minor amount of residual ESCRT-II in the siRNA KD study being sufficient for supporting functional budding. Notably in these studies, cell proliferation was largely unperturbed by the depletion of EAP45 and yet there was up to a 10-fold decrease in virus spread, suggesting that EAP45 is redundant for cell growth but critical for viral multiplication and spread. This strongly emphasises the importance of understanding the mechanistic roles of ESCRT-II in HIV budding.

Building on previous biochemical evidence [17, 18] from our collaborators in the Lever Group demonstrating the importance of ESCRT-II in efficient viral budding, we used TIRF microscopy [19] to visualise and quantify the spatial proximity of a SNAP-labelled EAP45

with Gag at the plasma membrane in fixed cells. We deleted parts of the interacting region between EAP45 and Gag and compared the co-occurrence of these functionally compromised mutants with that of the full length EAP45. Furthermore, live HeLa cells were also imaged using TIRF to visualise the dynamics of the interaction between EAP45 and Gag. Single particle tracking was used to obtain quantitative information from live cell movies, and the spatial proximity and dwell time of the particles were measured as indicators of interaction. The following sections describe the development of the imaging and analysis workflows used to measure colocalisation of EAP45 and Gag, the challenges encountered with different fluorescent labels, and the results obtained.

## **4.2 Imaging and analysis methods to study Gag and EAP45 at the plasma membrane**

To provide visual evidence of the reported interaction between EAP45 and Gag, two imaging strategies were implemented: First, *d*STORM was implemented with EAP45 and Gag labelled with commercial antibodies, however, this labelling strategy was shown to be highly unspecific, as the antibodies had not been previously validated. Conventional, diffraction-limited TIRF microscopy was employed instead using a SNAP-tag to label EAP45, and an endogenous GFP tag for Gag; both tags were compatible with live-cell imaging and were shown to be highly specific. A nearest-neighbour algorithm was developed for colocalisation analysis in fixed cells, along with a workflow for removing chromatic offset from the dual-colour images and segmenting EAP45 and Gag puncta from TIRF images of the plasma membrane. Finally, a single particle tracking workflow was developed for analysing the dynamic interaction between EAP45 and Gag in live cell movies.

### **4.2.1 *d*STORM imaging using conventional antibodies for Gag and EAP45**

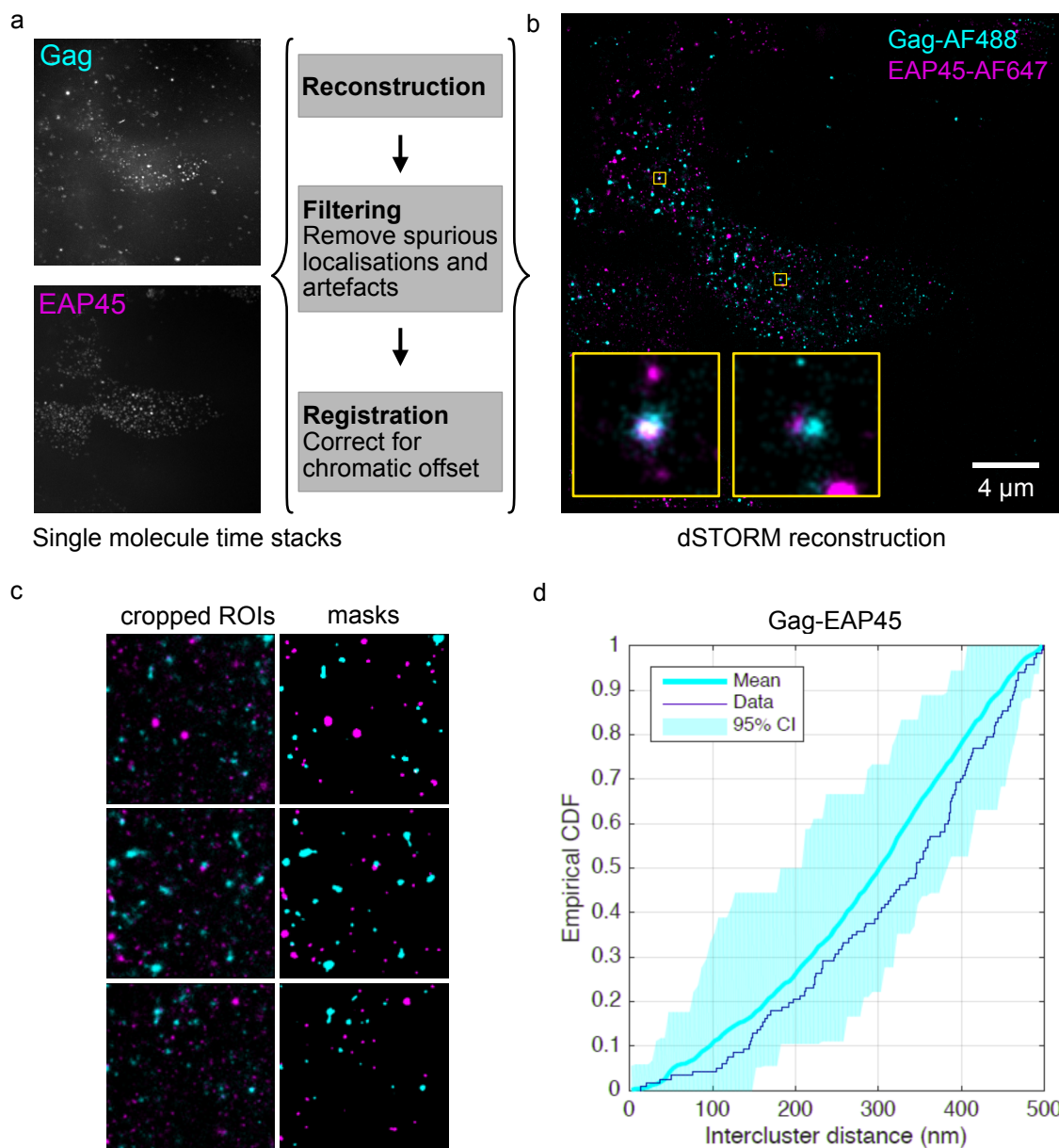
SMLM has been previously implemented to measure the proximity of Gag clusters to other ESCRT components [7, 8]. We implemented *d*STORM using the same microscope described in Section 3.2.3 of Chapter 3 to measure the inter-cluster distance between EAP45 and Gag puncta at the plasma membrane. HeLa cells were transfected with wild type HIV provirus, and subsequently fixed and immunostained using commercial antibodies for Gag (Anti-HIV-1 p24 Hybridoma 183-H12-5C, NIH AIDS Reagents) and EAP45 (HA Tag Antibody 14-6756-81, ThermoFisher). The antibodies were conjugated to an Alexa Fluor 647 (AF647) dye for EAP45, and an Alexa Fluor 488 (AF488) dye for Gag. These two dyes were chosen to avoid

spectral overlap. Bright and sparse puncta were visualised near the plasma membrane for both EAP45 and Gag upon widefield TIRF screening (4.2 a). Dr. Bo Meng performed all sample preparation steps (cell culture, transfection, and staining), and my contribution began when the samples were prepared and arrived at the microscope. All the imaging and analysis results shown this point onward were my responsibility.

*d*STORM image stacks were acquired sequentially for the 647 and 488 channels using a single camera, to avoid excitation of the red dye with the 488 nm laser. The time stacks obtained were reconstructed using thunderSTORM [21] and registered to remove chromatic offset using the same procedure as described in the Section 3.4.1 of Chapter 3 (Fig. 4.2). The resulting localisation data tables were filtered to remove localisations with an uncertainty larger than 40 nm, and intensities lower than 500 photons to avoid false positives. Two-colour composite images were generated from the reconstructed images, with multiple regions showing overlapping Gag and EAP45 clusters, as shown in the insets in Fig. 4.2 b.

The intercluster (centre-to-centre) distance between EAP45 and Gag clusters was calculated as a measure of colocalisation. First, a square region of interest (ROI) was chosen inside the cell to avoid detecting loose dye particles on the coverslip. The clusters in the cropped ROIs were segmented (Fig. 4.2 c) and their centroids were obtained for each channel. The centroid for every Gag cluster was used to calculate the intercluster distance to every EAP45 cluster. To compare the measured intercluster distance from the data to a random chance distribution of markers in the ROIs, the segmented clusters were randomly shuffled and their intercluster distances were measured in 1000 trials following Costes' method [20]. The cumulative distribution functions (CDF) of the intercluster distances from the experimental data and the random shuffling trials were plotted in Fig. 4.2 d. The CDF from the experimental data showed no significant difference with the CDF for a random chance distribution of markers. This experiment indicated that the measured proximity between Gag and EAP45 clusters was no different from the proximity from a simulated random arrangement of clusters.

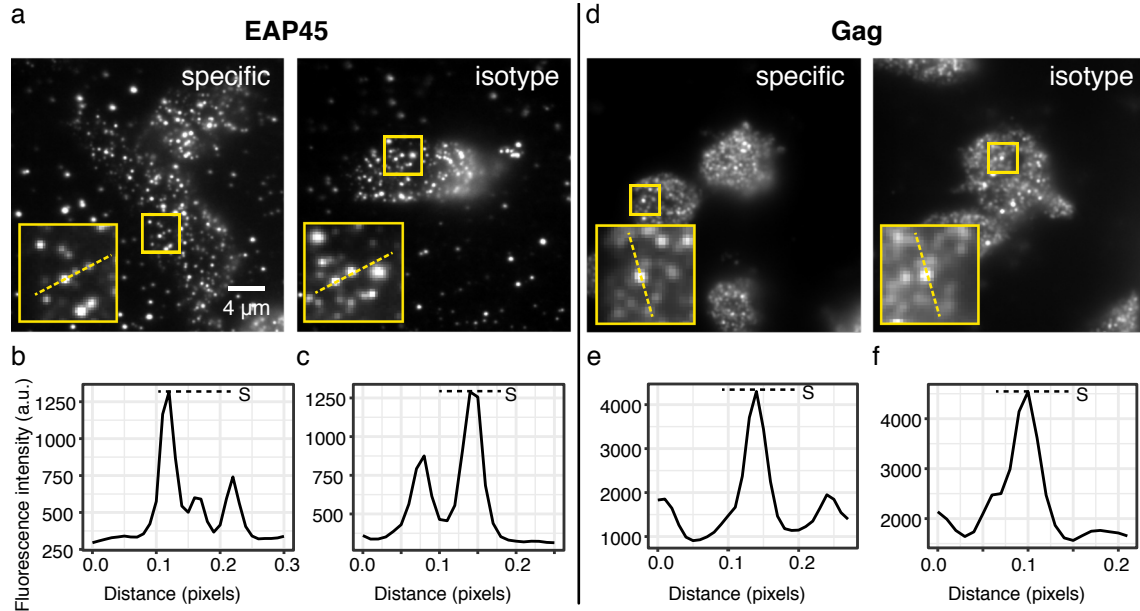
This result could have been due to a low number of true colocalisation events in the sample, or to the high density of clusters at the membrane leading to many overlapping events during the random shuffling process. However, many bright Gag and EAP45 puncta had been observed outside the cell boundaries during the imaging experiments, which were suspected to be non-specifically bound antibodies. To verify the specificity of the antibodies used, widefield TIRF was used to screen cells stained with isotype control antibodies for Gag and EAP45, which lack specificity to the target and therefore provide a reference for unspecific binding. Both specific and isotype control antibodies for EAP45 (Fig. 4.3 a-c) exhibited



**Fig. 4.2** *d*STORM imaging and analysis to measure the co-occurrence between Gag and EAP45. (a) *d*STORM image acquisition and reconstruction steps (b) Representative dual-colour super-resolution image of the cells stained for Gag (cyan) and EAP45 (magenta). (c) ROIs were selected inside the cells for intercluster distance analysis and Costes [20] shuffling (d) Cumulative distribution function of the likelihood of finding EAP45 clusters at a given distance from a Gag cluster, calculated with the results from both our experiments and the random chance simulations displayed with a 95% confidence interval.



puncta with similar intensities, suggesting there were a large number of fluorescently labelled false positives.



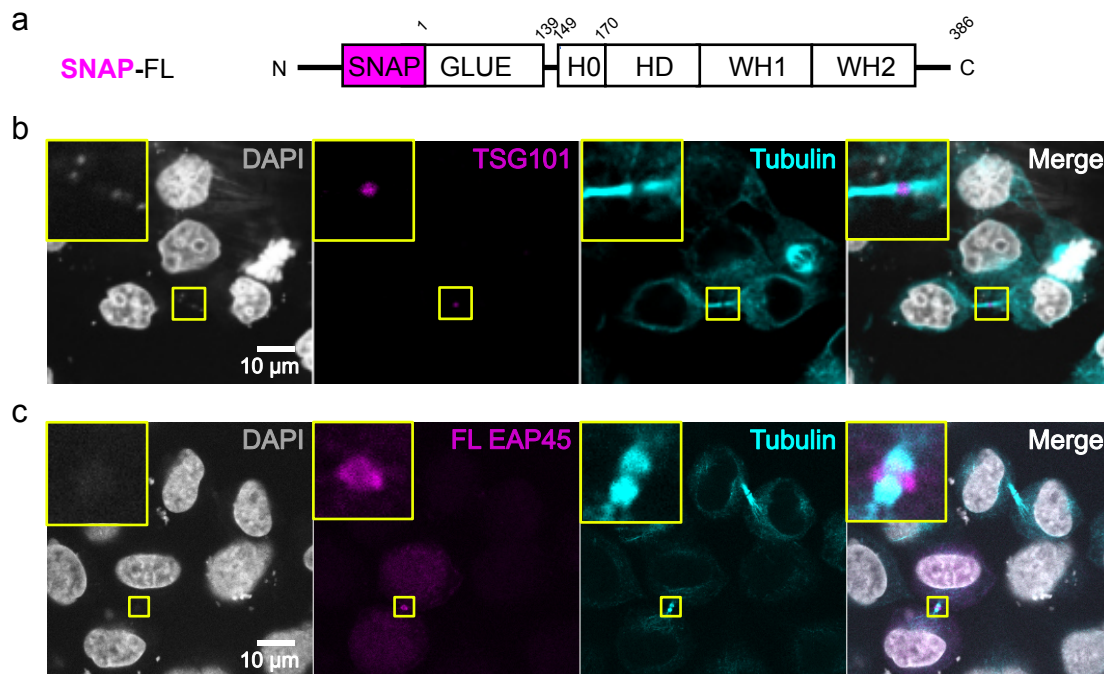
**Fig. 4.3** Images of HeLa cells immunostained for EAP45 (a) and Gag (b) with specific and isotype control antibodies. Yellow insets show line profiles traced for EAP45 specific (b) and isotype (c), and Gag specific (e) and isotype (f) puncta, to demonstrate the similarity in peak signal (S) levels for both antibody types.

This was also the case for the Gag antibodies (Fig. 4.3 d-f). To compare the fluorescent signal between specific and isotype control antibodies, line profiles were traced through bright features as shown in the yellow insets for Fig. 4.3 a-c (Eap45) and Fig. 4.3 d-f (Gag). These experiments confirmed the antibody specificity was low, and therefore the dyes were likely non-specifically bound to cellular structures other than our desired species.

#### 4.2.2 Validation of a SNAP-tag label for EAP45

Following the negative results from the *d*STORM experiments with conventional antibodies, more specific fluorescent labels were necessary, ideally compatible with live-cell imaging. A GFP construct previously validated for producing GFP-Gag intracellularly [22] was used to label Gag, and a SNAP-tag was used to label EAP45. The SNAP-tag is a polypeptide that can be fused to any protein of interest, and further conjugated with high specificity to a fluorescent dye. SNAP-tag (~20 kDa) provides similar molecular specificity to that of fluorescent proteins, and it is widely used for imaging both fixed and live cells due to the versatility of the commercially available substrates [23]. GFP, a slightly larger tag with ~27

kDa, had been previously inserted at the N-terminus of EAP45 without adversely affecting its cellular localisation [24], so we decided to SNAP label EAP45 at the same site (Fig. 4.4 a). The following section details the transfection of GFP and SNAP-tag for Gag and EAP45.



**Fig. 4.4** Verification of the SNAP EAP45 labelling in cytokinesis. (a) Schematic showing the full length (FL) EAP45 construct used in this study tagged with a SNAP-tag at the N-terminus. (b) Confocal images of HeLa cells expressing YFP-TSG101 stained with DAPI for nuclei and  $\alpha$ -tubulin show TSG101 is recruited to the intercellular bridge. (c) Images of HeLa cells transfected with SNAP-FL EAP45 followed by staining with DAPI,  $\alpha$ -tubulin and AF647-conjugated SNAP substrate also show EAP45 is recruited to the intercellular bridge during cytokinesis. The SNAP-tag immunostaining and confocal imaging experiments were performed by Dr. Bo Meng. This figure is courtesy of Dr. Bo Meng.

### 4.2.3 Transfection of GFP and SNAP-tag to label Gag and EAP45

HeLa cells were transfected with Fugene HD (E2311 Promega) and HAP1-EAP45 KO cells with Turbofectin (TF81001 Origene), as detailed by Meng et al. [17]. For the fixed cell imaging experiments, HeLa cells were co-transfected with 100 ng piGFP and 60 ng of SNAP-FL-EAP45. A HeLa cell line stably expressing mCherry-ALIX was transfected with 100 ng piGFP, for the ALIX experiments. HAP1-EAP45 KO cells were co-transfected with 200 ng piGFP and 72 ng of SNAP-FL-EAP45. At 24 hours post transfection, cells were fixed, permeabilised, and stained with 4 μM SNAP-Surface AlexaFluor647 (S9136S New England Biolabs) for 1 hour. For live cell experiments, 120 ng of piGFP and HVPgagpro-

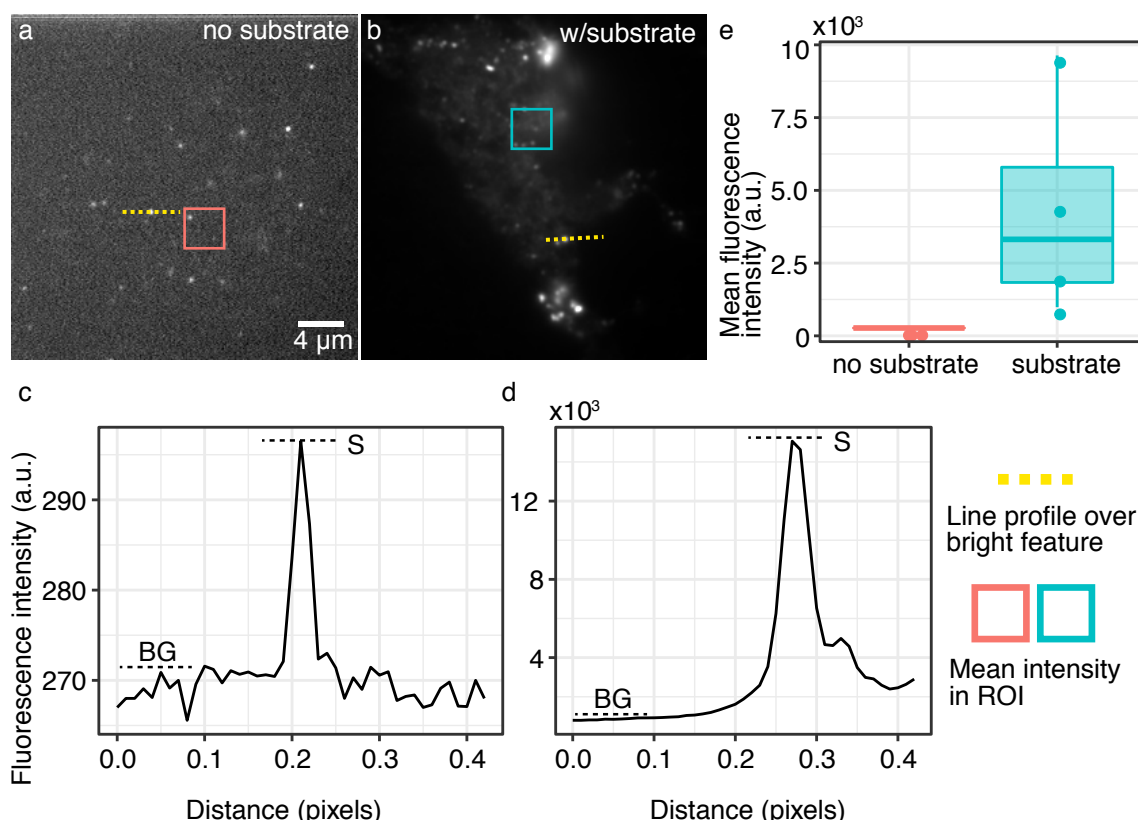
[17] plasmids at a 1:5 ratio were co-transfected with 72 ng of SNAP-FL-EAP45. At 10 hours post-transfection, cells were stained with 3  $\mu$ M cell-permeable live SNAP-cell 647 SiR (S9102S New England Biolabs) for 30 minutes prior to imaging. All cell culture, transfection, and staining were performed by Dr. Bo Meng.

### **Fusion of a SNAP-tag at the N-terminus of EAP45 does not affect its normal cellular localisation**

A full length (FL) SNAP-EAP45 was transfected into HeLa cells, and a confocal microscope was used to verify the recruitment of EAP45 to the intercellular bridge during cytokinesis. This is one of the activities well known to involve ESCRT-II, in cooperation with ESCRT-I and ESCRT-III [25, 26]. HeLa cells expressing YFP-TSG101, the ESCRT-I protein known to be recruited early in the membrane scission process, were used as a control [25]. Fig. 4.4 b shows TGS101 was recruited to the intercellular bridge, as expected. Following SNAP staining with an AF647-conjugated substrate, the SNAP-labelled EAP45 was also recruited as shown in Fig. 4.4 c. This suggests the normal cellular localisation of EAP45 is not affected by the insertion of an N-terminal SNAP-tag.

### **SNAP-tag is a highly specific label for EAP45**

To confirm the specificity of the SNAP-tag label, HeLa cells were imaged in TIRF illumination both with (Fig. 4.5 b) and without (Fig. 4.5 a) the SNAP substrate added. The images of the cells without substrate showed predominantly background signal close to the camera baseline ( $\sim 200$  counts) as shown in Fig. 4.5 a. The images of the cells with the substrate were bright, with a discernible cell outline and bright puncta near the plasma membrane (Fig. 4.5 b). Line profiles were traced over the brightest features of both the cells without substrate and with substrate, and the ratio of signal to background (SBR) was estimated for each using the peak value and the average background value from the line profiles. The cells with the SNAP substrate had an SBR approximately seven times higher than that of the cells without substrate. To account for differing background fluorescence levels, the mean intensity over a square ROI without any bright features shown with a red square in Fig. 4.5 a and a blue square in Fig. 4.5 b was plotted in Fig. 4.5 e. The mean intensity in the ROI of the cells with the SNAP substrate was over 4-fold higher than that of the cells with no substrate. Both of these measures indicate the SNAP-tag label yielded specific and bright labelling, suitable for testing colocalisation with GFP-Gag.

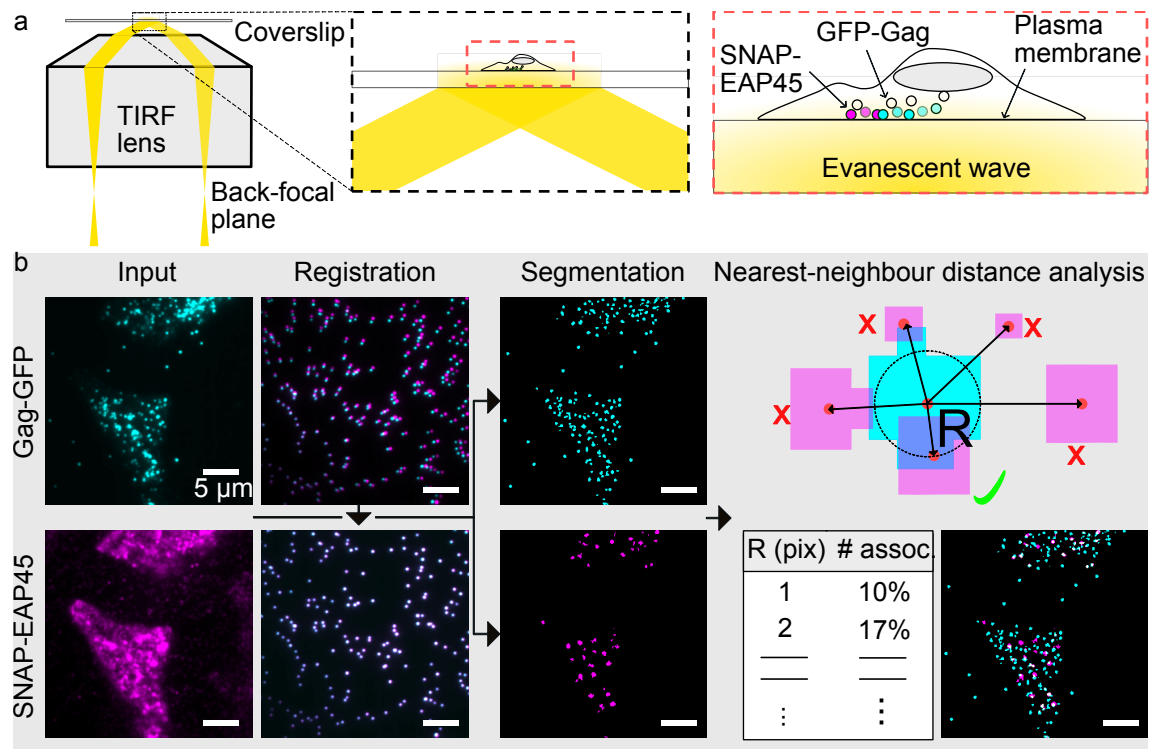


**Fig. 4.5** Verifying the specificity of the SNAP-tag label for EAP45. (a) Cells with no SNAP substrate added. A line profile was drawn across a bright feature and plotted in (c), showing the mean signal ( $\sim 270$  a.u.) is predominantly background, and the SBR is  $\sim 1$ . (b) Cells with SNAP substrate added. The mean signal from a line profile through a bright feature plotted in (d) is well over 1000 a.u., and the SBR is  $\sim 7$ . (e) Comparison of the mean signal inside a square ROI taken both in the case of no substrate and substrate. The contrast levels in images (a) and (b) are different to allow us to visualise the much lower intensity features in (a). Each data point in (e) represents the mean intensity computed over a region of interest in an image, for four total images acquired, each corresponding to a separate technical replicate.

#### 4.2.4 Nearest-neighbour distance analysis to measure spatial co-occurrence from TIRF images

Dual-colour diffraction limited images were acquired in TIRF on the same widefield microscope described in Chapter 3, this time using both cameras with a dichroic beam splitter (T565spxr-UF3) for simultaneous two-colour imaging. The co-occurrence between SNAP-EAP45 and GFP-Gag was measured using a nearest-neighbour distance analysis [27, 28] to search for EAP45 particles in the vicinity of Gag particles. Based on the average diameter of HIV budding virions ( $\sim 120$  nm) [29] and the diffraction-limited resolution of our microscope

(2 pixels=234 nm), a 1-pixel perimeter was selected as a constraint for defining potential interactions between the two species. Any Gag particles found to have a neighbouring EAP45 within this threshold were labelled as “associated”. To account for different densities of Gag particles in each cell, the number of associated particles was expressed as a percentage of the total number of detected particles. Fig. 4.6 a-b show a visual summary of the implemented imaging set up and workflow.



**Fig. 4.6** Imaging EAP45 and Gag at the plasma membrane using TIRF microscopy. (a) Illumination configuration used to image Gag and EAP45 at the plasma membrane in cells. (b). Workflow for evaluation the spatial proximity between EAP45 and Gag, including steps for chromatic offset correction using multi-fluorescent sub-diffractive beads, segmentation, and nearest-neighbour analysis.

In the workflow shown in Fig. 4.6 b, the raw dual-colour images were first registered to correct for chromatic offset using the DoM Utrecht plugin v1.1.6 [30] in Fiji [31], then segmented using the Trainable Weka Segmentation (TWS) plugin [32] from Fiji. For TWS, the first 3 images in each data set were manually annotated to identify three object classes: single fluorescent spots in focus (small blobs), large aggregates of spots (large blobs), and background; the Random Forest algorithm with 200 decision trees and 2 random features per set up was then used to classify the remaining images in the data set. Binary masks were

then created from the top 30th percentile of the class probability maps output by the TWS plugin for the small blob class. The large blob and background classes were discarded.

A nearest-neighbour algorithm was developed in Matlab from the ground up to check, for each segmented spot in the masks, the distance from its nearest neighbours in the opposite colour channel using the formula given by Eq. 2.20 from Chapter 2. Channel 1 spots found within the association distance threshold defined above of channel 2 spots were labelled as “associated”, and vice-versa. The algorithm developed can be summarised as follows:

1. Import binary masks for both colour channel 1 and channel 2.
2. For each mask, classify the segmented spots using the Matlab function *regionprops* to obtain their centroid.
  - (a) For each spot in channel 1, calculate its centre-to-centre distance to every spot in channel 2.
  - (b) Count the number of channel 1 spots which have at least one neighbour in channel 2 closer than  $R = 1$  pixel.
3. Repeat for all masks and export results in .csv file

In terms of computational intensity, the most time-consuming step in the colocalisation workflow described was the Trainable Weka Segmentation, as manual annotation for each data set could take up to 10 mins. All other steps could be completed within a few minutes, leading the entire workflow to employ less than an hour for completion. The source code for colocalisation analysis can be found in the repository for this project at [https://github.com/pedropabloVR/Meng\\_VallejoRamirez\\_EAP45\\_supportingCode](https://github.com/pedropabloVR/Meng_VallejoRamirez_EAP45_supportingCode).

#### 4.2.5 Single particle tracking of Gag and EAP45 particles in HeLa cells

Concurrent with the experiments to measure the spatial co-occurrence of Gag and EAP45 in fixed cells, we also sought to analyse the temporal interaction between these two proteins in live cells. The assembly of Gag particles at the cell membrane during HIV budding has been reported to occur within a time frame of ~10 mins [33, 34]. Subsequent recruitment of ESCRT proteins by Gag has been reported for TSG101 [34], VPS4A [35, 7, 4], some ESCRT-III proteins [4, 34], and ALIX [36], with an association time at the membrane varying from a few seconds up to ~12 minutes [37].

To visualise the dynamic association between GFP-Gag and SNAP-FL EAP45 at the plasma membrane, 20-min movies of live HeLa cells were captured using TIRF illumination.

A single particle tracking workflow was developed to analyse the amount of time EAP45 spends within the defined distance association threshold ( $\sim 117$  nm) to a Gag particle. First, dual-colour live cell movies were registered to correct for chromatic offset using the DoM Utrecht plugin [30] in Fiji.  $\sim 80 \times 80$  pixels<sup>2</sup> ( $\sim 1.3 \times 1.3$   $\mu\text{m}^2$ ) regions in which EAP45 particles were found in the vicinity of a Gag particle were then manually extracted. Gag and EAP45 particles were tracked separately using TrackMate v4.0.0 [38], which detects spots in each frame with sub-pixel precision, links their positions from frame to frame with Jaqaman's linear assignment problem (LAP) tracker [39], and outputs tracks. Using TrackMate, the tracks from a single movie could be generated within a few minutes, after optimising parameters for spot detection and trajectory linking using the graphical user interface. Finally, a custom Matlab script for nearest neighbour analysis was written from scratch to calculate the distances between the centre positions of the Gag and EAP45 tracks. The algorithm used for this script can be summarised as follows:

1. Import track files a 3-column array format, containing frame number and the x and y coordinates, for each colour channel.
2. Check for non-consecutive frames in each array, e.g. instances in which the tracked particle skips one or more frames. Fill in the frame gap using the particle's previous location.
3. Calculate the distance between the x,y centre coordinates of the particle in channel 1 to that in channel 2 using Eq. 2.20.
4. Store results in a .csv file for export.

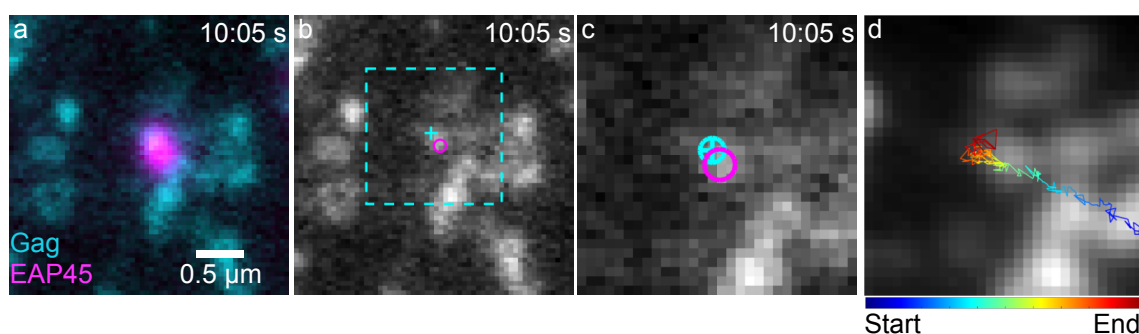
The detector used in the live cell experiments was a Hamamatsu Orca Flash v4.0 with a pixel pitch of 6.45  $\mu\text{m}$ , which translated into  $\sim 50$  nm pixels at the image plane after accounting for the objective (100x) and relay lens (1.3x) magnification. The association distance threshold used for the results from the live cell experiments was thus a minimum of 2 pixels. The source code for the track analysis can be found in the code repository for this project at [https://github.com/pedropabloVR/Meng\\_VallejoRamirez\\_EAP45\\_supportingCode](https://github.com/pedropabloVR/Meng_VallejoRamirez_EAP45_supportingCode).

#### **4.2.6 Co-moving frame analysis to visualise proximity between two moving particles**

In the live cell movies acquired, Gag and EAP45 particles move simultaneously, making it difficult to see how their positions are correlated over time. To visualise a correlated motion of the particles more clearly, an algorithm was developed in Matlab that tracked the position



of a Gag particle and repositioned the field of view such that the detected particle remained in the centre of the image from frame to frame. We used the Gag particle position as a reference point against which the relative motion of the EAP45 particle could be visualised. We call this algorithm a “co-moving frame analysis”, and its steps are summarised in Fig. 4.7. First, a region of interest (ROI) is chosen in the live cell movie (Fig. 4.7 a), and the centroid of both the Gag particle (cyan cross) and the EAP45 particle (magenta circle) are detected using a sub-pixel fitting routine from TrackMate [38], as shown in Fig. 4.7 b. The output tracks are imported into Matlab, and a square region around the Gag particle is cropped and extracted, such that the Gag particle is always in the center (dotted cyan rectangle, Fig. 4.7 b). The position of the EAP45 particle is then plotted relative to the new position of the Gag particle (Fig. 4.7 c and d). The result is the track of the EAP45 particle motion, colour-coded for time, plotted over the average image of the immobilised Gag particle in the center of the field of view (Fig. 4.7 d). This helps visualise the dwell times of the moving particle around the fixed particle.



**Fig. 4.7** Co-moving frame analysis to analyse the temporal interaction between Gag and EAP45 (a) Dual colour frame from a live cell movie showing a single EAP45 particle moving in the vicinity of a Gag particle. (b) The centre position of the tracks for Gag (cross) and EAP45 (circle) were overlaid on a grey scale frame of the Gag channel. A dotted rectangle indicates a region around the centre position of the Gag particle which is extracted as a new frame of reference. (c) The position of the EAP45 molecule is then plotted relative to the new frame of reference. (d) Track of the EAP45 particle motion colour-coded from start to finish of the track duration, overlaid on an average image of the Gag channel.

#### 4.2.7 Statistical analysis

For the case in which only two experimental data sets were compared, such as in Fig. 4.9 for ALIX and FL EAP45, and Fig. 4.12 for FL EAP45 and  $\Delta G\Delta H$ , Welch's two sample t-test was used with ns: non-significance, \* for  $p < 0.05$ , \*\* for  $p < 0.01$ , and \*\*\* for  $p < 0.001$ . Boxplots were overlaid on the raw data points, with the bold horizontal line showing the



mean, and whiskers represent the largest and smallest value within 1.5 times the interquartile range above 75th and 25th percentile, respectively. For the case in which more than two conditions were compared (Fig. 4.11), a One-Way ANOVA with Dunnett Comparison was used to compare the FL EAP45 condition to each of the functionally compromised mutants ( $\Delta G\Delta H$ ,  $\Delta PTAP$ ,  $\Delta PTAP\text{-}\Delta G\Delta H$ ) with ns: non-significance, \* for  $p < 0.05$ , \*\* for  $p < 0.01$ , and \*\*\* for  $p < 0.001$ . At least three biological replicates were tested for all experimental conditions imaged, and the number  $n$  reported for the association of Gag and EAP45 represents the number of cells in which colocalisation was evaluated. All statistical tests were performed using R using ggplot2 [40]. The code and the data files to replicate all figures and analysis performed can be found in the code repository for this project at [https://github.com/pedropabloVR/Meng\\_VallejoRamirez\\_EAP45\\_supportingCode](https://github.com/pedropabloVR/Meng_VallejoRamirez_EAP45_supportingCode).

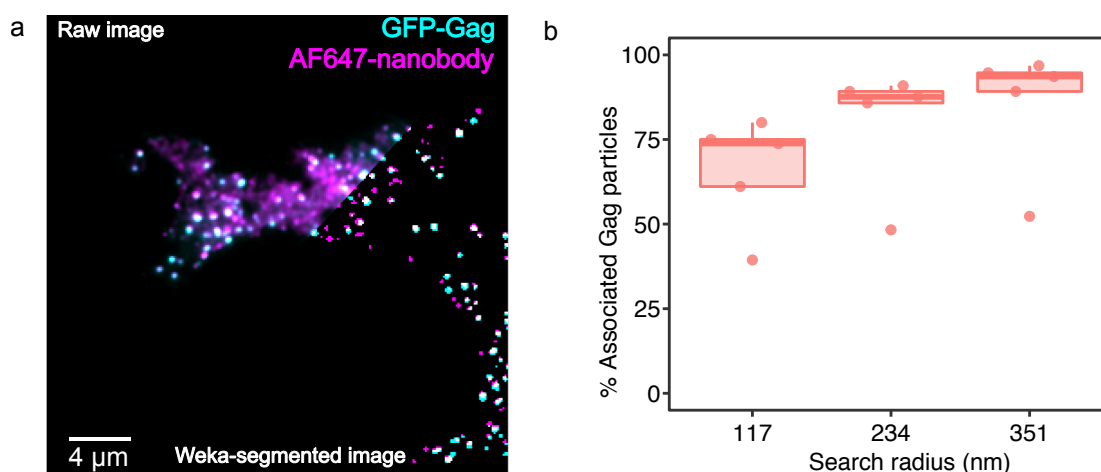
## 4.3 Results

The nearest-neighbour analysis to study colocalisation in fixed cells described above was validated using two control experiments: First, an anti-GFP fluorescent nanobody was used to test the maximal experimentally attainable co-occurrence with GFP-Gag; Next, the co-occurrence of Gag and the ALIX protein, which is known to be recruited to the budding site during viral egress and to colocalise with Gag [36], was compared to that of Gag and FL EAP45. Having validated the nearest-neighbour analysis, it was applied to measure the co-occurrence of Gag and EAP45 mutants with a functionally compromised N-terminus. In addition, the temporal interaction between Gag and FL EAP45 was examined using the single particle tracking and co-moving frame analysis workflow.

### 4.3.1 Testing maximal experimentally available colocalisation using anti GFP nanobody with Alexa Fluor 647 dye

GFP-Gag was immunostained with an anti-GFP nanobody labelled with AF647 in HeLa cells. The nanobody binds to GFP with a small linker length (~2 nm) such that all the GFP fluorescent spots would also have an AF647 spot in practically the same location. A representative dual-colour image of the GFP-Gag (cyan) and the AF647-nanobody (magenta) is shown in Fig. 4.8 a, with a blended-in image of the segmented mask towards the bottom right corner. The results from the nearest-neighbour analysis are shown in Fig. 4.8 b, with the percentage of associated GFP-Gag particles plotted as a function of search radius, in pixel-wise increments (1 pixel = 117 nm). A maximum search radius of 3 pixels was used, as particles farther apart than 3 pixels were unlikely to be associated. Despite expecting every

GFP-Gag detected to have an attached nanobody, nanobody binding is not 100% efficient and only ~75% of the detected GFP-Gag particles have a nanobody within 1 pixel (117 nm), and ~87% within 2 pixels (234 nm). This discrepancy may in addition be due to an error in the centroid estimation after segmentation, introducing a bias into the nearest-neighbour analysis and causing partially overlapping particles to be not counted as “associated”.



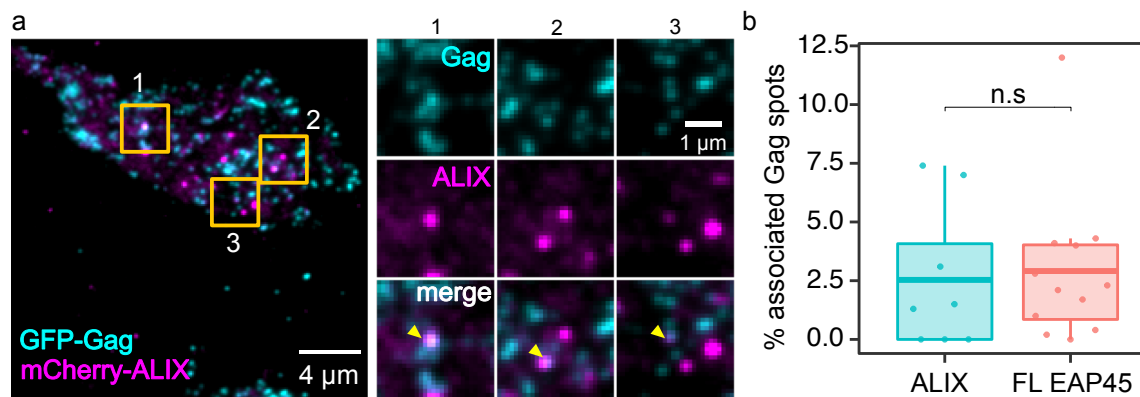
**Fig. 4.8** Using an anti-GFP AF647 nanobody to test the maximal experimentally attainable colocalisation with GFP-Gag. (a) Representative dual-colour image of GFP-Gag and the AF647-nanobody, with a faded-in composite towards the lower right corner showing the resulting dual-colour mask after segmentation. (b) Percentage of associated Gag particles from nearest-neighbour analysis. Each data point corresponds to the measured associated Gag particles for a single cell. A total of five cells from three different replicates were analysed.

### 4.3.2 FL EAP45 and ALIX show similar co-occurrence with Gag

A HeLa cell line expressing mCherry-ALIX was transfected with GFP-Gag to test the co-occurrence between these two proteins. ALIX is known to be transiently recruited to the budding site by Gag for subsequent recruitment of other ESCRT components during viral egress [36]. The colocalisation of Gag assembly sites and ALIX clusters has been previously reported by other groups, both as a centre-to-centre distance measurement using dSTORM ( $89 \pm 61$  nm) [7], and as a percentage of Gag assembly sites with a neighbouring ALIX over the total number of Gag assembly sites (3.4%) [8]. Despite a very low expression of the mCherry-ALIX, i.e. there were very few cells which had bright ALIX puncta which we could analyse, the images for 8 cells were analysed. On average, ~2.5% of the total Gag particles had an ALIX particle within a 1-pixel distance (Fig. 4.9 b). This colocalisation

rate is similar to the 1-4% reported previously by Prescher et al. for ALIX and other ESCRT components [8].

The colocalisation rate between Gag and ALIX was compared to that between Gag and FL EAP45 in a HeLa cell line transfected with SNAP-EAP45 stained with an AF647 dye. After analysing 12 cells, ~3% of the total Gag particles had a FL EAP45 particle within a 1-pixel distance, which is similar to the association rate observed for Gag and ALIX. This similarity suggests that in HeLa cells, both EAP45 and ALIX might be recruited to the budding sites close to the plasma membrane in a similar abundance. The results from the co-occurrence analysis are shown in Fig. 4.9 b. A representative cell with GFP-Gag and mCherry-ALIX is shown in Fig. 4.9 a, with yellow-boxed insets highlighting overlapping Gag and ALIX particles.

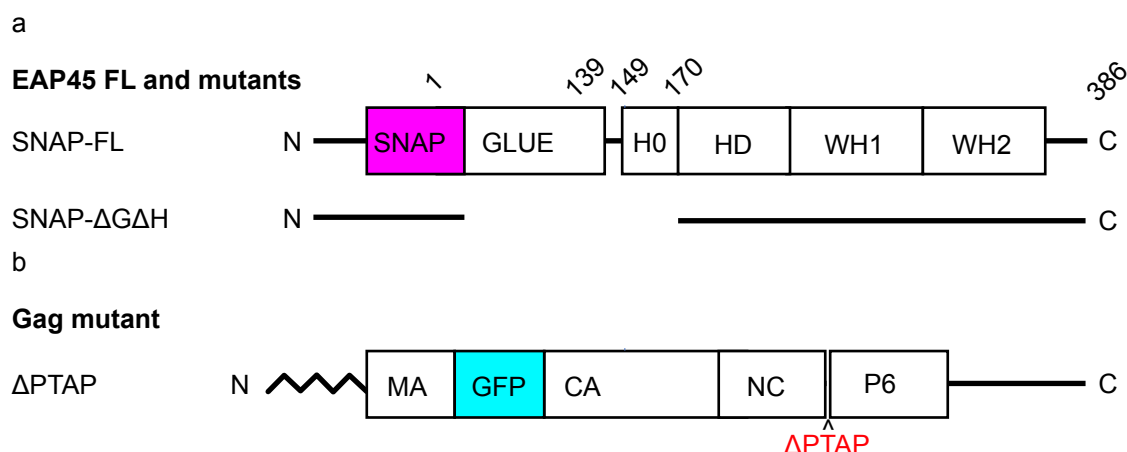


**Fig. 4.9** Co-occurrence analysis between ALIX/Gag and SNAP-FL EAP45/Gag. (a) HeLa mCherry-ALIX (magenta) cells transfected with GFP-Gag (cyan). Yellow insets highlight regions with visible co-occurrence of ALIX and Gag spots. (b) Co-occurrence analysis of Gag/ALIX compared to that of Gag/FL EAP45 (separate HeLa cells were co-transfected with SNAP-FL EAP45 and GFP). Each data point represents a cell ( $n = 8$  cells for ALIX, and  $n = 12$  cells for FL EAP45) processed with the pipeline illustrated in Fig. 4.6.

### 4.3.3 FL EAP45 shows a higher co-occurrence with Gag than functionally compromised EAP45 mutants

Having observed a similar co-occurrence of Gag with ALIX and FL EAP45 (Fig. 4.9 b), we examined how deleting a part of the linker region between Gag and EAP45 would affect their co-occurrence. The Glue and H0 domains of EAP45 have been shown to interact with ESCRT-I [10]; their deletion ( $\Delta G\Delta H$ , Fig. 4.10 a), has been demonstrated to impair efficient viral budding in experiments using biochemical and virological methods [18].

To probe the effect of truncated Glue and H0 domains, we imaged HeLa cells transfected with GFP-Gag and SNAP-FL EAP45 (Fig. 4.11 a), and cells transfected with GFP-Gag and

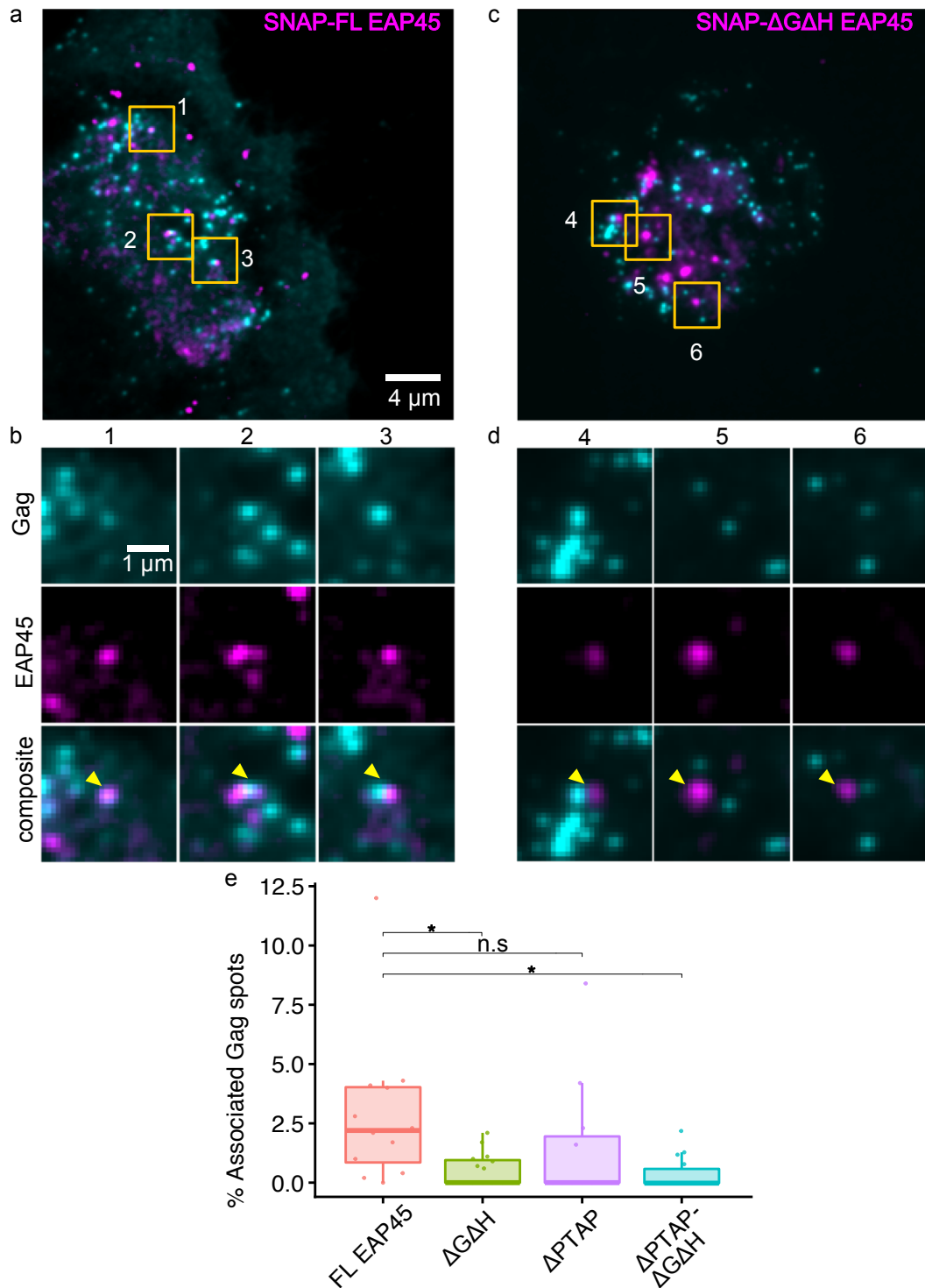


**Fig. 4.10** Functionally compromised mutants of EAP45 and Gag. (a) Diagram of the FL EAP45 and the construct with a deleted Glue and H0 domain, ΔGΔH-EAP45. (b) Diagram of the Gag polyprotein with a deleted PTAP domain. These diagrams are courtesy of Dr. Bo Meng.

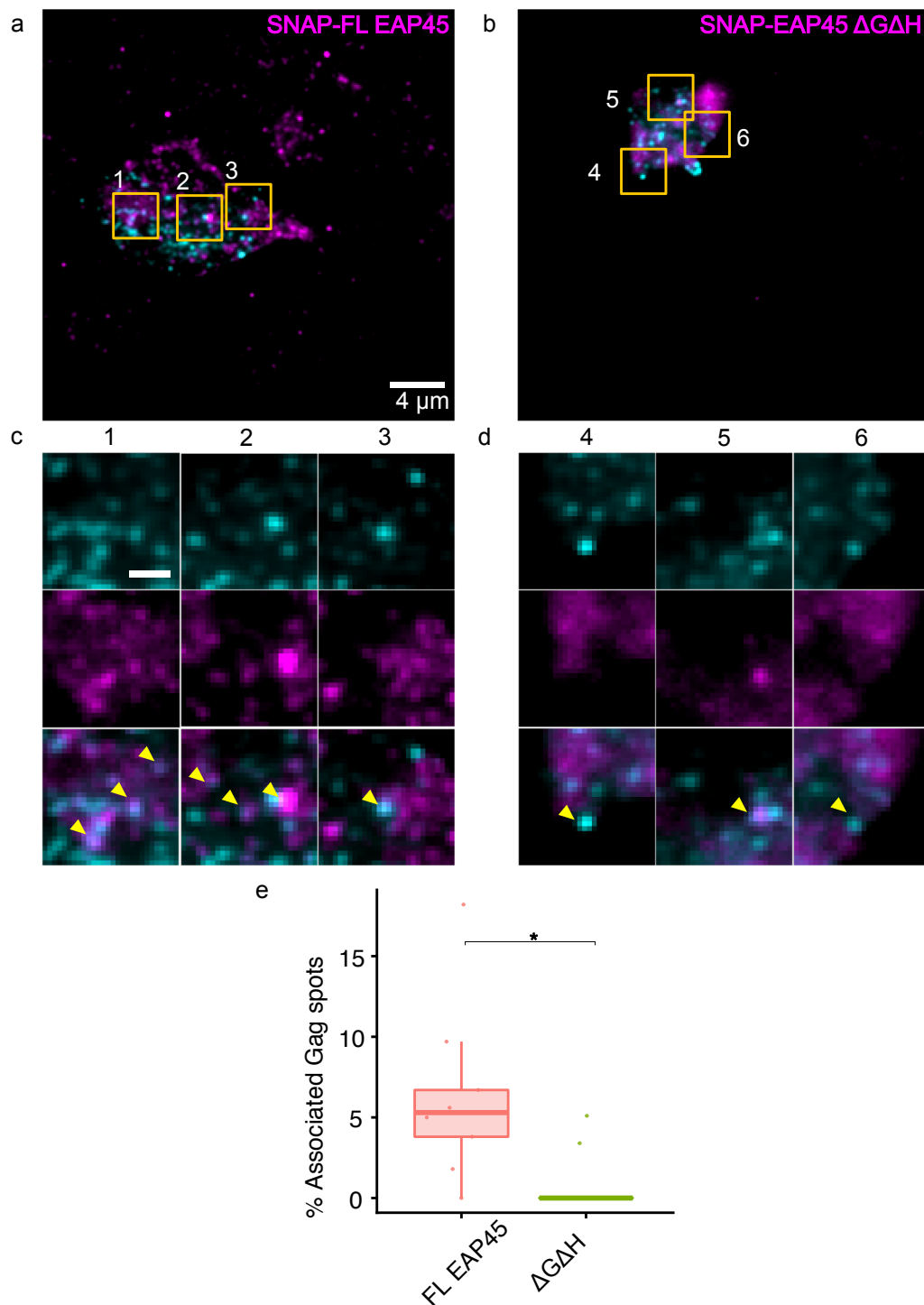
SNAP-ΔGΔH EAP45 (Fig. 4.11 b). Both were stained with an AF647-conjugated SNAP substrate. The nearest-neighbour analysis workflow showed the spatial proximity between the ΔGΔH EAP45 and Gag (n=15 cells) was lower (Fig. 4.11 e) than that measured for the FL EAP45 and Gag (n= 12 cells,  $p = 0.026$  between FL EAP45 and ΔGΔH EAP45).

The PTAP domain in Gag (Fig. 4.10 b) is involved in recruiting TSG101 (ESCRT-I), likely initiating the recruitment cascade of ESCRT-II and -III in HIV budding [6]. We hypothesised the removal of the PTAP domain (ΔPTAP) would decrease the co-occurrence of Gag and EAP45. Our data however show that even though there is a decrease in the average co-occurrence between ΔPTAP Gag (n = 11 cells) and FL EAP45, this was not significantly different from that of wildtype Gag and FL EAP45 (Fig. 4.11 e,  $p = 0.36$ ). This suggests EAP45 co-occurs with Gag independently of PTAP. This observed co-occurrence with ΔPTAP relies on an intact Glue and H0 linker region, as the removal of this region (ΔPTAP + ΔGΔH) significantly reduces the observed co-occurrence (n = 14 cells, Fig. 4.11 e,  $p = 0.018$  between FL EAP45 and ΔPTAP ΔGΔH).

To further explore the effect observed, the co-occurrence between EAP45 and Gag was also examined in HAP1-EAP45 KO cells, with supplementation *in trans* of FL EAP45. The HAP1-EAP45 KO cell line was developed in a previous study by Meng et al. [17, 18], and used to demonstrate the arrest of HIV budding when all endogenous EAP45 is deleted using CRISPR-CAS9 [41, 42]. The nearest-neighbour analysis workflow showed the spatial proximity between Gag and FL EAP45 (n = 8 cells, Fig. 4.11 a and c) was significantly higher ( $p = 0.018$ ) than that between Gag and ΔGΔH EAP45 (n = 9 cells, Fig. 4.11 b and d).



**Fig. 4.11** Co-occurrence of GFP-Gag and different functionally compromised SNAP-EAP45 mutants in fixed HeLa cells. Representative TIRF images for cells expressing FL EAP45 (a) with numbered insets showing regions of high spatial overlap between markers (c), and  $\Delta\text{G}\Delta\text{H}$  EAP45 (b) with numbered insets showing regions with little spatial overlap between markers (d). (e) Comparison of the percentage of associated Gag particles between FL EAP45 and  $\Delta\text{G}\Delta\text{H}$  EAP45 expressing HeLa cells, as well as those with  $\Delta\text{PTAP}$ -Gag mutants. Each data point corresponds to the measured associated Gag particles for a single cell. Three different biological replicates were analysed for each condition.



**Fig. 4.12** Co-occurrence of GFP-Gag and SNAP-EAP45 in fixed HAP1 EAP45 KO cells. Representative TIRF images for cells expressing FL EAP45 (a) with numbered insets showing regions of high spatial overlap between the two markers (c), and  $\Delta\text{G}\Delta\text{H}$  EAP45 (b) with numbered insets showing regions with little spatial overlap between markers (d). (e) Comparison of the percentage of associated Gag particles, i.e. those which have an EAP45 particle within 1 pixel, between FL EAP45 and  $\Delta\text{G}\Delta\text{H}$  EAP45 expressing HAP1 EAP45 KO cells. Each data point corresponds to the measured associated Gag particles for a single cell. Three different biological replicates were analysed for each condition.

The findings from the nearest-neighbour analysis in HeLa and HAP1 EAP45 KO cells suggest the Glue domain and the H0 linker region play an important role in linking EAP45 to HIV budding [17, 18].

#### **4.3.4 Single particle tracking of live HeLa cells shows transient co-occurrence of Gag and EAP45**

The recruitment of ESCRTs by Gag at the plasma membrane to achieve viral egress from the cell is a sequential process lasting up to 30 minutes [37]. Individual recruitment events for ESCRT-I and III proteins have been reported to last between a few seconds and 12 minutes [35, 7, 4, 34, 36]. To visualise the temporal interaction between Gag and ESCRT-II, live HeLa cells were transfected with GFP-Gag, similarly to previous reports imaging ESCRT component recruitment by Gag in live cells [7, 36], and with SNAP-FL EAP45. The SNAP-EAP45 substrate was added at >10 hours post-transfection, when enough Gag molecules had assembled at the plasma membrane to be detected as bright puncta. 20-min movies with frames every 5 to 10 seconds were recorded, and 14 regions of interest were extracted and analysed using the single particle tracking and co-moving frame analysis methods described in the Methods section. Three representative regions from the live cell recordings are shown in Fig. 4.13 a at different time points, along with the results from the co-moving frame analysis in Fig. 4.13 b. The extracted ROIs represent areas in the videos in which EAP45 was observed to be in close proximity to Gag and were chosen manually. For each ROI, the distance between the two tracks was plotted as a function of time (Fig. 4.13 b). The number of consecutive frames in which EAP45 particles were found to stay within the association threshold of Gag were counted, resulting in dwell times ranging from 10 seconds to 6 mins, as shown in the kernel density distribution in Fig. 4.13 c). From this distribution, the mean consecutive association time between EAP45 and Gag was estimated to be ~2 mins. This co-occurrence time frame agrees with that reported for other ESCRT proteins [34]. Within all the 14 analysed ROIs (Fig. 4.14) there appear to be three classes of movements observed between EAP45 and Gag:

- EAP45 approaches Gag within the two-pixel threshold multiple times (Fig. 4.14 a; 7 out of 14 events).
- EAP45 approaches Gag within the two-pixel threshold and moves away (Fig. 4.14 b; 4 out of 14 events).
- EAP45 approaches Gag and remains within the two-pixel threshold during the time of the recording (Fig. 4.14 c ; 3 out of 14 events).

The percentage of the total observation time the EAP45 and Gag particles spent within a given distance of each other was plotted in Fig. 4.13 d). For all regions analysed, EAP45 particles spend ~30% percent of the total observation time within the diffraction-limited association threshold from a Gag particle.

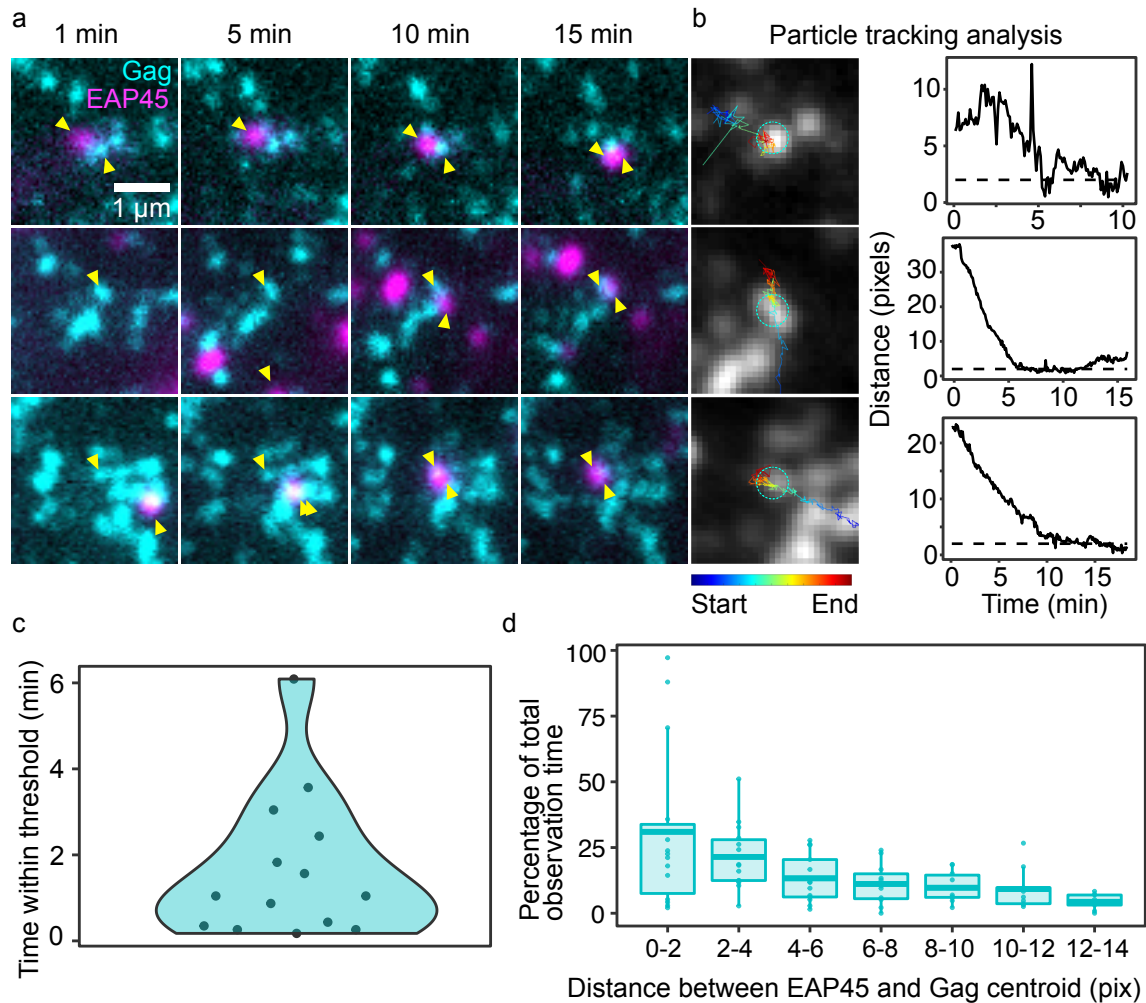
## 4.4 Summary and Discussion

The canonical view on membrane scission via the ESCRT complexes during HIV budding considered ESCRT-II to be unnecessary [11]. However, recent studies of ESCRT-II in cytokinesis [14, 43] and HIV budding [17, 18] revise this view and implicate ESCRT-II in a bridging role between ESCRT-I and ESCRT-III, as originally reported in cargo sorting [44, 9]. Previous biochemical analyses by our collaborators in the Lever group reported significantly impaired HIV budding and egress in HAP1 cells in which EAP45 had been knocked out using CRISPR-Cas9 [41, 42]. By expressing EAP45 *in trans*, efficient HIV budding and egress could be recovered. This effect was attributed to the H0 linker domain in EAP45 [17]. Building on their biochemical work, this chapter provides new evidence on the colocalisation of ESCRT-II with Gag during viral budding using TIRF microscopy. More specifically, this work sheds light on the domain requirements of the linker region between EAP45 and Gag for efficient viral budding, as well as the interaction dynamics between these two proteins.

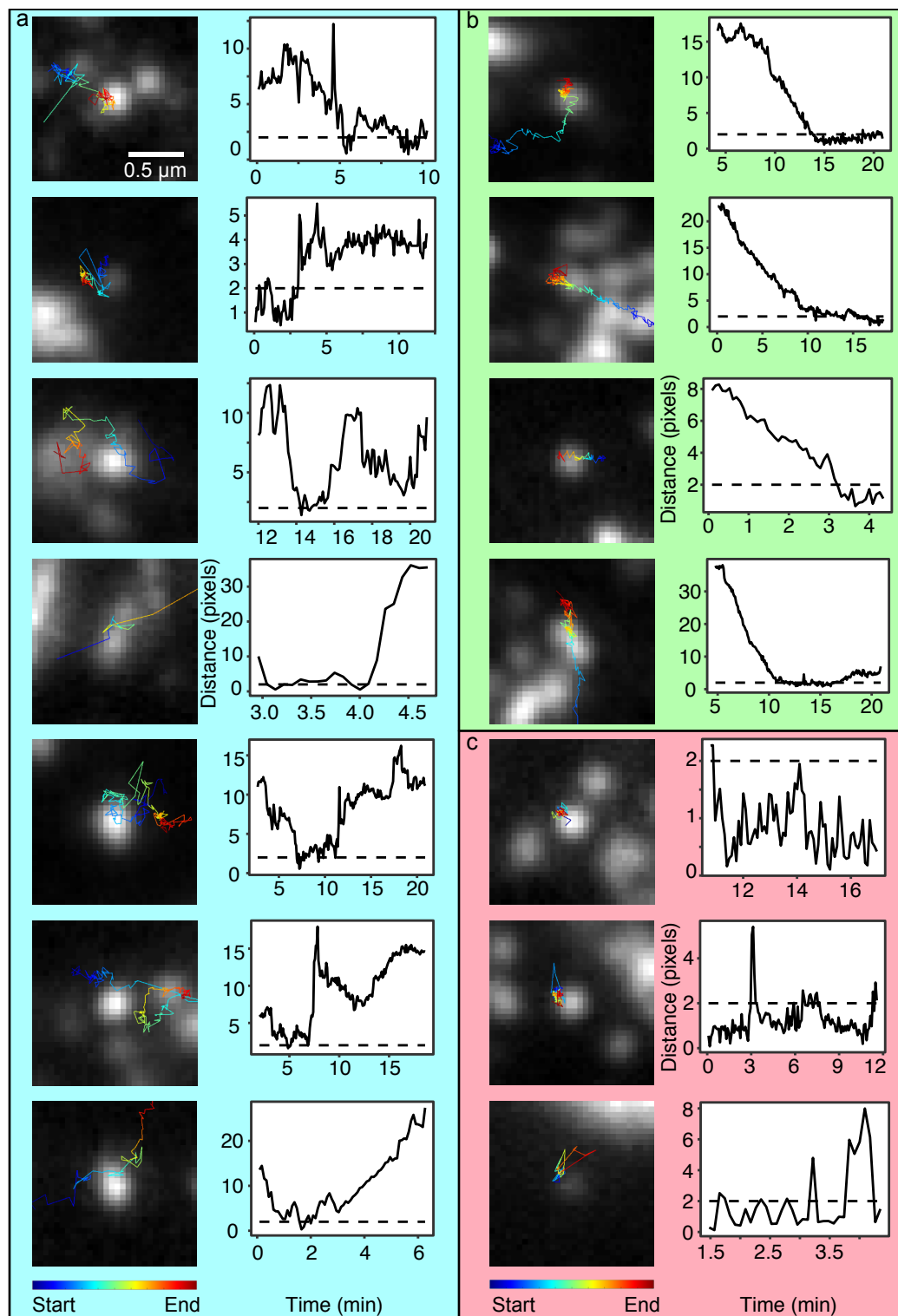
### 4.4.1 SNAP-tag and a GFP fusion provide higher specificity than conventional antibodies to fluorescently label EAP45 and Gag, respectively

The experiments using *d*STORM produced co-occurrence results between Gag and EAP45 which were no different from those obtained from a random distribution of the two markers; this was likely due to the use of highly unspecific antibodies during immunostaining. Despite looking like promising samples with the right shape, brightness, and distribution at the plasma membrane, the images with specific antibodies were shown to be no different from those of isotype control antibodies. As an alternative, a SNAP-tag label was used for EAP45 and a GFP construct to label Gag, both providing high molecular specificity and brightness in fixed and live cells. Our collaborators in the Lever lab verified the SNAP-tag would not affect the normal cellular location of the EAP45 by visualising its recruitment to the intercellular bridge during cytokinesis (Fig. 4.4 b and c). Because the GFP-Gag was not compatible with *d*STORM, diffraction-limited widefield imaging with TIRF illumination was implemented





**Fig. 4.13** Co-moving frame analysis on single particle tracking data shows transient co-occurrence of EAP45 and Gag. (a) Four time points for different representative ROIs. Yellow arrowheads point out Gag and EAP45 spots selected for tracking. (b) Results from the co-moving frame analysis for the three regions in (a). The path traced by the EAP45 particle is color-coded for time, overlaid on an average image of the stabilised Gag channel. Plots of the distance in pixels between the centroids of the Gag and EAP45 particles versus time are shown, with the dotted line representing the association threshold of 2 pixels (~100 nm). (c) Kernel density distribution showing the maximum consecutive association time between Gag and EAP45 over 14 ROIs. (d) Percentage of total observation time as a function of the distance between centroids of Gag and EAP45, plotted over 2-pixel increments (n = 14). Each data point in (d) represents an analysed region in which a single EAP45 and a Gag particle were tracked, for a total of 14 regions analysed.



**Fig. 4.14** Co-moving frame analysis reveals three classes of EAP45 movement relative to Gag: (a) EAP45 particles which hover in the vicinity of the Gag cluster for short spans of time, and either oscillate in and out of proximity, or move away from the Gag cluster, (b) EAP45 particles which approach the virus from a distance and remain in the 2-pixel vicinity of the virus particle for several minutes, and (c) EAP45 particles which remain within 2 pixels of the virus particle throughout the observation time interval.

for the rest of the experiments. The loss in imaging resolution and sensitivity was traded off for the high specificity provided by the new labels.

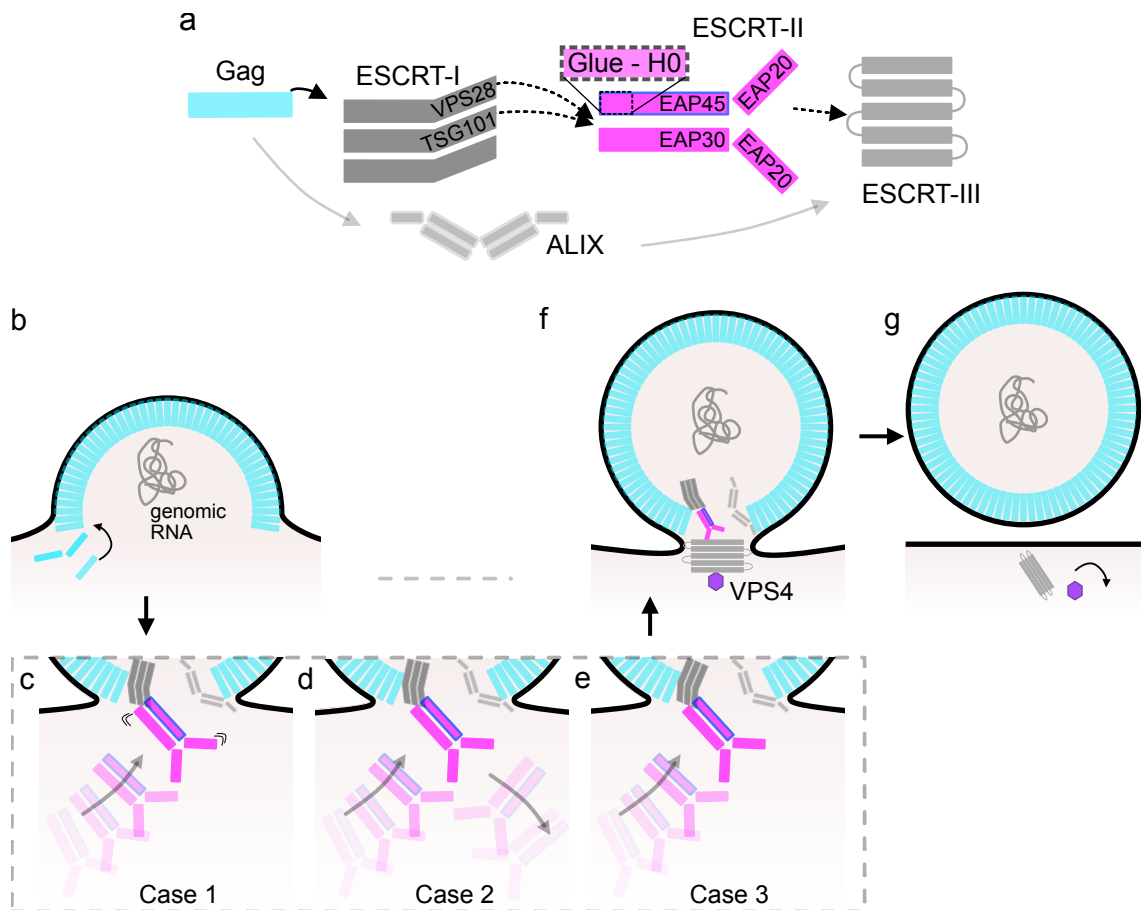
#### 4.4.2 The EAP45 N-terminus is necessary for colocalisation with Gag

The co-occurrence of Gag and FL EAP45 was significantly higher than that of Gag and EAP45 constructs with a deleted Glue and H0 domain ( $\Delta$ G $\Delta$ H Fig. 4.11 e and Fig. 4.12 e), confirming the N terminal region of EAP45 is important in interacting with Gag at the plasma membrane. A higher co-occurrence rate between FL EAP45 and Gag was measured in HAP1-EAP45 KO cells compared to that in HeLa cells (Fig. 4.12 e and Fig. 4.11 e). This could be explained by EAP45's necessary involvement in HIV budding in HAP1 EAP45 KO cells, in which only externally supplemented SNAP-EAP45 is present. In HeLa cells, endogenously expressed EAP45, which is unlabelled and therefore cannot be detected, could also function *de novo* thereby explaining the lower co-occurrence rate measured in HeLa cells.

The co-occurrence of Gag and EAP45 appears to have a weak dependence on the PTAP motif in Gag, as its deletion does not significantly reduce the measured co-occurrence compared to wild type Gag (Fig. 4.11 e). Given that the level of recruitment and dynamics at the plasma membrane between wild type and  $\Delta$ PTAP Gag are similar [35, 34, 45], this weak dependence could be explained by the recruitment of EAP45 by ALIX via its known interaction with the ESCRT-I TSG101 protein [46] as a compensatory pathway in the absence of the PTAP motif.

Altogether, the results from the nearest-neighbour analysis provide the first visual evidence of the spatial association between ESCRT-II and Gag, which strengthens previous work showing the H0 region in EAP45 is critical for successful viral budding [17] via its link to ESCRT-I [12, 11]. Fig. 4.15 a shows ESCRT-II's proposed position in the sequential ESCRT recruitment sequence by Gag first presented in Fig. 4.1.

The low co-occurrence rate (~1 to 5%) measured between FL EAP45 and Gag (Fig. 4.11 e and Fig. 4.12 e) and ALIX and Gag (Fig. 4.9 c) is similar to the previously reported co-occurrence rates of other ESCRT components and Gag (1.5 to 3.4%) using *d*STORM [8]. The low co-occurrence observed in our study may to some extent be due to errors in the centroid estimation after segmentation of the particles. For instance, in regions where particles were too densely clustered, the centroid of a particle could be wrongly estimated due to the presence of bright neighbouring particles. This could introduce a bias into the nearest-neighbour analysis, causing partially overlapping particles to be not counted as "associated". The low number of cells in the ALIX experiments ( $n = 8$ ) was due to a very low incidence of bright cells in the mCherry-ALIX cell line; it was very difficult to find



**Fig. 4.15** Proposed models for the role of ESCRT-II in HIV budding (a) Current understanding of the ESCRT protein apparatus in HIV budding, highlighting the N-terminus of EAP45 as a proposed linker to Gag/ESCRT-I (b) Gag accumulation causes deformation of the plasma membrane, leading to the recruitment of ESCRTs. Three possible models of ESCRT-II recruitment are illustrated: (c) recruitment in which the ESCRT-II oscillates towards and away from the bud, (d) transient recruitment followed by decoupling, and (e) persistent recruitment. (f) Final stage before membrane scission in which the polymerised ESCRT-III is recruited to the budding neck and acts in combination with VPS4 to achieve membrane scission and bud egress (g).

cells which expressed both mCherry-ALIX and GFP-Gag and had bright puncta in both colour-channels. Similarly, the low fluorescent signal from SNAP-EAP45 and GFP-Gag in both the HeLa and HAP1 cells made it difficult to find cells which had bright, discernible signal in both colour channels, resulting in a low number of cells for colocalisation analysis.

#### **4.4.3 Three potential dynamic interactions between EAP45 and Gag were observed**

The measured proximity between EAP45 and Gag in the live imaging experiments could be indicative of a dynamic interaction between these two proteins at viral assembling sites (Fig. 4.13 and 4.14). Since the excitation region due to the TIRF illumination was confined to a thin volume above the coverslip, and the illumination intensity decreases exponentially outside this region, any motion of the cells or drifting of the sample holder in a direction perpendicular to the imaging plane had a large effect on the intensity measurements of the Gag and EAP45 puncta at the membrane. The focus over time was maintained manually, therefore the position of the focal plane was not maintained deterministically and the intensity of the puncta fluctuated due to the focusing motion. Therefore, the changes in the intensity of the puncta over time [7] were not considered a reliable measure of accumulation or recruitment. Instead, the centre-to-centre distances of Gag and EAP45 puncta were calculated and plotted (Fig. 4.13) to visualise how the motion of EAP45 changed relative to its proximity to Gag as way to infer a potential interaction.

For 30% of the total recording time in all ROIs, EAP45 was located within ~100 nm of Gag. Three potential classes of EAP45 movement relative to Gag were observed (Fig. 4.14 a-c). Based on these preliminary data sets, the possible cases for ESCRT-II's dynamic interaction with Gag were illustrated in Fig. 4.15 c-e. However, these data are not sufficient to predict or infer any mechanistic function of EAP45 during HIV budding. Future studies are necessary to pinpoint which class of dynamic interaction(s) is necessary and responsible for forming of infectious virus particles as well as the recruitment kinetics of ESCRT- II in the context of other ESCRT proteins.

### **4.5 Future work**

Additional studies are necessary to determine the exact interactions of ESCRT-II with ESCRT-I and -III, and find out the mechanism of sequential recruitment by which ESCRT-II contributes to viral egress. A sensible follow-up experiment to the work in this Chapter could explore the spatial co-occurrence of TSG101/EAP45 using the same imaging and

analysis tools presented here, to corroborate previous biochemical evidence showing the link between ESCRT-I and -II [12, 11]. In addition, a multi-colour live cell imaging study with labels for Gag, TSG101, EAP45, and an ESCRT-III protein could be performed to record the sequential recruitment of Gag and all ESCRTs to the budding site. This would require a TIRF microscope with an autofocus system to maintain focus at the membrane, and a scanning stage to image many cells and increase the likelihood of observing ESCRT recruitment events. Another option would be to use a light-sheet microscope to observe HIV egress at the membrane in 3D, similarly to the study by Scherer et al. [47] in which light-sheet is used to examine different stages of herpes simplex virus 1 (HSV-1) replication. The gentle illumination, large field of view, and optical sectioning capabilities of this imaging modality could be very beneficial to study the spatiotemporal association of EAP45 with Gag, TSG101, and ESCRT-III proteins.

To improve on the sensitivity and accuracy of the colocalisation results presented in this chapter, dual-colour *d*STORM could be readily performed using the SNAP-AF647 EAP45, and an Alexa Fluor 568-conjugated anti-GFP nanobody to tag GFP-Gag. This would allow us to precisely measure the average interaction distance between EAP45 and Gag using the *d*STORM method described in section 4.2.1, adding to the library of known Gag-ESCRT interaction distances [7].

To further explore the effect of deleting parts of the EAP45 N-terminus on HIV egress, the number of successfully detached viral particles from live infected cells could be quantified. This could be accomplished by imaging Gag assembly sites with TIRF illumination in live cells, and artificially detaching the cells using Trypsin > 10 hours post-transfection to only detect released virus particles. A similar experiment has been reported by Bendjennat and Saffarian [48] to count virus-like particles attached to the coverslip as a measure of viral release. Alternatively, a fluorogenic dye which becomes bright only upon contact with a coverslip coating with a specific pH could be used to label Gag. This would cause only mature virions which exit the cell to be detected by TIRF illumination, allowing us to quantify the number of maturing virions in real time.

The challenges in understanding the role of ESCRT-II in HIV budding are numerous. Concerning the biological work, this involves culturing cells with the necessary constructs for Gag and ESCRT members, and the reliable fluorescent labelling of the species to be imaged, among other challenges. Concerning the imaging work, this involves high-throughput imaging using multiple colour channels both at the diffraction limit and using super resolution, as well as reliable image analysis methods to avoid false positives during registration, segmentation, and the co-occurrence analysis. Nonetheless, understanding the mechanistic

roles of ESCRT-II in HIV budding is important, as further insights into these protein-protein interactions could help in developing a potential new class of antivirals to treat HIV.

## 4.6 Conclusion

To conclude, we used a SNAP-tag label to visualise the spatial and temporal co-occurrence of the ESCRT-II protein EAP45 with the HIV polyprotein Gag in HIV budding. A nearest-neighbour analysis was developed to study the spatial co-occurrence of Gag and EAP45, and how it was affected by deleting parts of the putative interacting region between them. Deletion of the Glue and the H0 domains significantly reduced the co-occurrence between Gag and EAP45, highlighting the importance of this region in linking to Gag potentially via its link to ESCRT-I. Single particle tracking was applied to study the temporal co-occurrence of Gag and EAP45, and a co-moving frame analysis was newly developed to visualise the simultaneous movement of two particles more easily. EAP45 particles were found to dwell within 100 nm of Gag for as little as 10 seconds and up to 6 mins, suggesting a potential transitory association between Gag and EAP45 during viral egress. This study presented the first visual evidence of the association between ESCRT-II and Gag, and it paves the way for further experiments to investigate the mechanistic role of ESCRT-II in HIV budding.

## 4.7 Acknowledgment of contributions

The project described in this chapter was a collaborative effort between the Lever Lab in the Department of Medicine (University of Cambridge), and the Laser Analytics Group, involving the expertise of virologists, biochemists, and microscopists. The biological question driving the project was conceived by Prof. Andrew Lever, Dr. Julia Kenyon, and Dr. Bo Meng from the Lever Group. Dr. Bo Meng performed all the cell culture work, the fluorescent labelling and biochemistry verification of the labels for Gag and the ESCRT proteins, and the confocal imaging of the cytokinetic spindle. Imaging experiments (*d*STORM, widefield TIRF for fixed and live cells) were performed in collaboration with Dr. Bo Meng and Dr. Katharina Scherer. The code for colocalisation analysis for the *d*STORM experiments was written in collaboration between myself and Ezra Bruggeman. I wrote all other pieces of code, including the R scripts to plot analysis results and to run statistical analysis and the code for the co-moving frame analysis. I performed the image analysis for the widefield TIRF experiments with the GFP-Gag, the SNAP-EAP45, and the functionally compromised mutants as well as all the analysis for the live cell imaging data. I would like to thank Prof. Clemens Kaminski, Dr. Eric J Rees, Prof. Andrew Lever, and Dr. Julia Kenyon for their

valuable questions, comments, and suggestions regarding the imaging and analysis work in this project.

## References

- [1] Bo Meng, Pedro P Vallejo Ramirez, Katharina M Scherer, Ezra Bruggeman, Julia C Kenyon, Clemens F Kaminski, and Andrew M Lever. Distinct domain requirements for EAP45 in HIV budding, late endosomal recruitment, and cytokinesis. *bioRxiv*, page 2020.05.23.112607, jan 2020.
- [2] Roger L. Williams and Sylvie Urbé. The emerging shape of the ESCRT machinery. *Nature Reviews Molecular Cell Biology*, 8(5):355–368, 2007.
- [3] Jörg Votteler and Wesley I. Sundquist. Virus budding and the ESCRT pathway. *Cell Host and Microbe*, 14(3):232–241, 2013.
- [4] Daniel S. Johnson, Marina Bleck, and Sanford M. Simon. Timing of ESCRT-III protein recruitment and membrane scission during HIV-1 assembly. *eLife*, 7:1–20, 2018.
- [5] Bo Meng and Andrew M.L. Lever. Wrapping up the bad news - HIV assembly and release. *Retrovirology*, 10(1):1, 2013.
- [6] Wesley I Sundquist and Hans-Georg Krausslich. HIV-1 Assembly, Budding, and Maturation. *Cold Spring Harbor Perspectives in Medicine*, pages 1–24, 2012.
- [7] Marina Bleck, Michelle S. Itano, Daniel. S. Johnson, V. Kaye Thomas, Alison J. North, Paul D. Bieniasz, and Sanford M. Simon. Temporal and spatial organization of ESCRT protein recruitment during HIV-1 budding. *Proceedings of the National Academy of Sciences*, 111(33):12211–12216, 2014.
- [8] Jens Prescher, Viola Baumgärtel, Sergey Ivanchenko, Adriano A Torrano, Christoph Bräuchle, Barbara Müller, and Don C. Lamb. Super-Resolution Imaging of ESCRT-Proteins at HIV-1 Assembly Sites. *PLoS Pathogens*, 11(2):1–26, 2015.
- [9] Hsiangling Teo, Olga Perisic, Beatriz González, and Roger L. Williams. ESCRT-II, an endosome-associated complex required for protein sorting: Crystal structure and interactions with ESCRT-III and membranes. *Developmental Cell*, 7(4):559–569, 2004.
- [10] David J. Gill, Hsiangling Teo, Ji Sun, Olga Perisic, Dmitry B. Veprintsev, Scott D. Emr, and Roger L. Williams. Structural insight into the ESCRT-I/-II link and its role in MVB trafficking. *EMBO Journal*, 26(2):600–612, 2007.



- [11] Charles Langelier, Uta K. von Schwedler, Robert D. Fisher, Ivana De Domenico, Paul L. White, Christopher P. Hill, Jerry Kaplan, Diane Ward, and Wesley I. Sundquist. Human ESCRT-II Complex and Its Role in Human Immunodeficiency Virus Type 1 Release. *Journal of Virology*, 80(19):9465–9480, 2006.
- [12] Young Jun Im and James H. Hurley. Integrated Structural Model and Membrane Targeting Mechanism of the Human ESCRT-II Complex. *Developmental Cell*, 14(6):902–913, 2008.
- [13] Shrawan Kumar Mageswaran, Natalie K. Johnson, Greg Odorizzi, and Markus Babst. Constitutively active ESCRT-II suppresses the MVB-sorting phenotype of ESCRT-0 and ESCRT-I mutants. *Molecular Biology of the Cell*, 26(3):554–568, 2015.
- [14] Liliane Christ, Eva M. Wenzel, Knut Liestøl, Camilla Raiborg, Coen Campsteijn, and Harald Stenmark. ALIX and ESC RT-I/II function as parallel ESC RT-III recruiters in cytokinetic abscission. *Journal of Cell Biology*, 212(5):499–513, 2016.
- [15] Ai Hui Tang, Haiwen Chen, Tuo P. Li, Sarah R. Metzbower, Harold D. MacGillavry, and Thomas A. Blanpied. A trans-synaptic nanocolumn aligns neurotransmitter release to receptors. *Nature*, 536(7615):210–214, 2016.
- [16] F. Ann Ran, Patrick D. Hsu, Jason Wright, Vineeta Agarwala, David A. Scott, and Feng Zhang. Genome engineering using the CRISPR-Cas9 system. *Nature Protocols*, 8(11):2281–2308, 2013.
- [17] Bo Meng, Natasha C.Y. Ip, Truus E.M. Abbink, Julia C. Kenyon, and Andrew M.L. Lever. ESCRT-II functions by linking to ESCRT-I in human immunodeficiency virus-1 budding. *Cellular Microbiology*, (June 2019):1–15, 2020.
- [18] Bo Meng, Natasha C.Y. Ip, Liam J. Prestwood, Truus E.M. Abbink, and Andrew M.L. Lever. Evidence that the endosomal sorting complex required for transport-II (ESCRT-II) is required for efficient human immunodeficiency virus-1 (HIV-1) production. *Retrovirology*, 12(1):1–15, 2015.
- [19] Daniel Axelrod. Total Internal Reflection Fluorescence Microscopy. *Traffic*, 2:764–774, 2001.
- [20] Sylvain V. Costes, Dirk Daelemans, Edward H. Cho, Zachary Dobbin, George Pavlakis, and Stephen Lockett. Automatic and quantitative measurement of protein-protein colocalization in live cells. *Biophysical Journal*, 86(6):3993–4003, 2004.

- [21] Martin Ovesný, Pavel Křížek, Josef Borkovec, Zdeněk Švindrych, and Guy M. Hagen. ThunderSTORM: A comprehensive ImageJ plug-in for PALM and STORM data analysis and super-resolution imaging. *Bioinformatics*, 30(16):2389–2390, 2014.
- [22] Wolfgang Hubner, Ping Chen, Armando Del Portillo, Yuxin Liu, Ronald E. Gordon, and Benjamin K. Chen. Sequence of Human Immunodeficiency Virus Type 1 (HIV-1) Gag Localization and Oligomerization Monitored with Live Confocal Imaging of a Replication-Competent, Fluorescently Tagged HIV-1. *Journal of Virology*, 81(22):12596–12607, 2007.
- [23] Antje Keppler, Maik Kindermann, Susanne Gendreizig, Horst Pick, Horst Vogel, and Kai Johnsson. Labeling of fusion proteins of O6-alkylguanine-DNA alkyltransferase with small molecules in vivo and in vitro. *Methods*, 32(4):437–444, 2004.
- [24] Thomas Slagsvold, Rein Aasland, Satoshi Hirano, Kristi G. Bache, Camilla Raiborg, Daniel Trambaiolo, Soichi Wakatsuki, and Harald Stenmark. Eap45 in mammalian ESCRT-II binds ubiquitin via a phosphoinositide- interacting GLUE domain. *Journal of Biological Chemistry*, 280(20):19600–19606, 2005.
- [25] Jez G. Carlton and Juan Martin-Serrano. Parallels between cytokinesis and retroviral budding: A role for the ESCRT machinery. *Science*, 316(5833):1908–1912, 2007.
- [26] Eiji Morita, Virginie Sandrin, Hyo Young Chung, Scott G. Morham, Steven P. Gygi, Christopher K. Rodesch, and Wesley I. Sundquist. Human ESCRT and ALIX proteins interact with proteins of the midbody and function in cytokinesis. *EMBO Journal*, 26(19):4215–4227, 2007.
- [27] E. Lachmanovich, D. E. Shvartsman, Y. Malka, C. Botvin, Y. I. Henis, and A. M. Weiss. Co-localization analysis of complex formation among membrane proteins by computerized fluorescence microscopy: Application to immunofluorescence co-patching studies. *Journal of Microscopy*, 212(2):122–131, 2003.
- [28] Thibault Lagache, Nathalie Sauvonnnet, Lydia Danglot, and Jean Christophe Olivo-Marin. Statistical analysis of molecule colocalization in bioimaging. *Cytometry Part A*, 87(6):568–579, 2015.
- [29] John A.G. Briggs, Kay Grünewald, Bärbel Glass, Friedrich Förster, Hans Georg Kräuslich, and Stephen D. Fuller. The mechanism of HIV-1 core assembly: Insights from three-dimensional reconstructions of authentic virions. *Structure*, 14(1):15–20, 2006.

- [30] Eugene A. Katrukha. DoM Utrecht plugins. [https://github.com/ekatrukha/DoM\\_Utrecht](https://github.com/ekatrukha/DoM_Utrecht), 2017.
- [31] Johannes Schindelin, Ignacio Arganda-Carreras, Erwin Frise, Verena Kaynig, Mark Longair, Tobias Pietzsch, Stephan Preibisch, Curtis Rueden, Stephan Saalfeld, Benjamin Schmid, Jean-Yves Tinevez, Daniel James White, Volker Hartenstein, Kevin Eliceiri, Pavel Tomancak, and Albert Cardona. Fiji: an open-source platform for biological-image analysis. *Nat Meth*, 9(7):676–682, jul 2012.
- [32] Ignacio Arganda-Carreras, Verena Kaynig, Curtis Rueden, Kevin W Eliceiri, Johannes Schindelin, Albert Cardona, and H Sebastian Seung. Trainable Weka Segmentation : a machine learning tool for microscopy pixel classification. *Bioinformatics*, 33(March):2424–2426, 2017.
- [33] Nolwenn Jouvenet, Paul D Bieniasz, and Sanford M Simon. Imaging the biogenesis of individual HIV-1 virions in live cells. *Nature*, 454(7201):236–240, 2008.
- [34] Nolwenn Jouvenet, Maria Zhadina, Paul D Bieniasz, and Sanford M Simon. Dynamics of ESCRT protein recruitment during retroviral assembly. *Nature Cell Biology*, 13(4):394–401, 2011.
- [35] Viola Baumgärtel, Sergey Ivanchenko, Aurélie Dupont, Mikhail Sergeev, Paul W Wiseman, Hans Georg Kräusslich, Christoph Bräuchle, Barbara Müller, and Don C Lamb. Live-cell visualization of dynamics of HIV budding site interactions with an ESCRT component. *Nature Cell Biology*, 13(4):469–476, 2011.
- [36] Pei I. Ku, Mourad Bendjennat, Jeff Ballew, Michael B. Landesman, and Saveez Saffarian. ALIX is recruited temporarily into HIV-1 budding sites at the end of Gag assembly. *PLoS ONE*, 9(5):1–8, 2014.
- [37] Nadia Ruthardt, Don C. Lamb, and Christoph Bräuchle. Single-particle tracking as a quantitative microscopy-based approach to unravel cell entry mechanisms of viruses and pharmaceutical nanoparticles. *Molecular Therapy*, 19(7):1199–1211, 2011.
- [38] Jean Yves Tinevez, Nick Perry, Johannes Schindelin, Genevieve M. Hoopes, Gregory D. Reynolds, Emmanuel Laplantine, Sebastian Y. Bednarek, Spencer L. Shorte, and Kevin W. Eliceiri. TrackMate: An open and extensible platform for single-particle tracking. *Methods*, 115:80–90, 2017.

- [39] Khuloud Jaqaman, Dinah Loerke, Marcel Mettlen, Hirotaka Kuwata, Sergio Grinstein, Sandra L. Schmid, and Gaudenz Danuser. Robust single-particle tracking in live-cell time-lapse sequences. *Nature Methods*, 5(8):695–702, 2008.
- [40] Hadley Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016.
- [41] Le Cong, F Ann Ran, David Cox, Shuailiang Lin, Robert Barretto, Naomi Habib, Patrick D Hsu, Xuebing Wu, Wenyan Jiang, Luciano A Marraffini, and Feng Zhang. Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science*, 339(6121):819 LP – 823, feb 2013.
- [42] Martin Jinek, Krzysztof Chylinski, Ines Fonfara, Michael Hauer, Jennifer A Doudna, and Emmanuelle Charpentier. A Programmable Dual-RNA – Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*, 337:816–822, 2012.
- [43] Inna Goliand, Dikla Nachmias, Ofir Gershony, and Natalie Elia. Inhibition of ESCRT-II-CHMP6 interactions impedes cytokinetic abscission and leads to cell death. *Molecular Biology of the Cell*, 25(23):3740–3748, 2014.
- [44] Hsiangling Teo, David J. Gill, Ji Sun, Olga Perisic, Dmitry B. Veprintsev, Yvonne Vallis, Scott D. Emr, and Roger L. Williams. ESCRT-I Core and ESCRT-II GLUE Domain Structures Reveal Role for GLUE in Linking to ESCRT-I and Membranes. *Cell*, 125(1):99–111, 2006.
- [45] Sebla B. Kutluay, Trinity Zang, Daniel Blanco-Melo, Chelsea Powell, David Jannain, Manel Errando, and Paul D. Bieniasz. Global changes in the RNA binding specificity of HIV-1 gag regulate virion genesis. *Cell*, 159(5):1096–1109, 2014.
- [46] Xi Zhou, Jiali Si, Joe Corvera, Gary E. Gallick, and Jian Kuang. Decoding the intrinsic mechanism that prohibits ALIX interaction with ESCRT and viral proteins. *Biochemical Journal*, 432(3):525–534, 2010.
- [47] Katharina M Scherer, James D Manton, Timothy K Soh, Luca Mascheroni, Viv Connor, Colin M Crump, and Clemens F Kaminski. Spatiotemporal analysis of host cell modification during herpes simplex virus 1 replication. *bioRxiv*, 2019.
- [48] Mourad Bendjennat and Saveez Saffarian. The race against protease activation defines the role of escrts in hiv budding. *PLOS Pathogens*, 12(6):1–25, 06 2016.

## **Part II**

### **Imaging at the mesoscale**



## Chapter 5

# Introduction to Optical Projection Tomography

The first part of this thesis described the implementation of single-molecule imaging and analysis methods in *in vitro* models of Parkinson's Disease and HIV. The second part of this thesis describes the development of an open-source, low-cost toolkit with which to image samples at the opposite end of the size spectrum, namely millimetre-scale tissues. The first chapter introduces the volumetric imaging technique applied in this toolkit, Optical Projection Tomography (OPT), to image millimetre-scale objects. A general introduction to volumetric imaging methods and the principles of OPT imaging and reconstruction are presented, along with previous biological applications and recent technology developments in OPT. The second chapter describes the development of OptiJ, an OPT system with off-the-shelf hardware for image acquisition and open-source software for reconstruction. OptiJ was applied to image dissected mouse lung samples in a proof-of-concept chronic obstructive pulmonary disease (COPD) study in collaboration with AstraZeneca.

### 5.1 Volumetric imaging techniques for biological tissue

The three-dimensional imaging of anatomical and functional features in mesoscopic (milli- to centimetre dimensions) samples provides valuable data for biomedical research. Common uses of 3D imaging in biomedical research include detecting abnormalities such as tumours in body parts and tissue samples, as well as recognizing disease phenotypes in model organisms.

### 5.1.1 Biomedical tomographic imaging techniques

In medical and clinical applications, whole body 3D imaging techniques such as computed tomography (CT) [1], which generates 3D reconstructions from 2D images acquired using x-rays, and magnetic resonance imaging (MRI) [2], which applies strong magnetic fields to map the location of water and lipids in tissue, are ubiquitous. CT and MRI are routinely used to visualise morphology in large tissues and organs on the scale of tens of centimetres. Both of these techniques are non-invasive, non-destructive, and highly effective at measuring morphological characteristics such as volume, shape, and density of opaque tissue, e.g. skin, muscle, and bone, in live samples. Micro-CT [3] can achieve micrometre-scale isotropic resolution [4] using contrast agents [5]. However, the benchtop instrumentation is very expensive (> £100,000) [6] and the high doses of radiation used make this unsuitable for repeated experiments on a biological sample, e.g. in longitudinal studies of embryonic development. Micro-MRI can also achieve micrometre-scale isotropic resolution [7] with excellent tissue contrast and no ionizing radiation. Similarly, the cost and size of MRI instruments can be prohibitive for many applications in research labs [7].

### 5.1.2 3D optical imaging techniques

Radiation in the visible spectrum (400-700 nm) has also been used extensively to image tissue either *in vivo* or *ex vivo* for medical diagnosis or biological research, however it has historically been limited to small penetration depths on the order of 10  $\mu\text{m}$ . When light travels through tissue, or any opaque medium, it is attenuated due to both absorption and scattering. Absorption occurs when the wavelength, or frequency of incoming photons matches the frequency associated with the medium molecules' energy transitions, and the photon energy is absorbed by the medium molecules. Absorption can either result in the emission of a lower frequency photon, or in energy dissipation in the form of heat, thereby reducing the number of available photons. Scattering is an elastic interaction between photons and biological tissue in which a photon is absorbed and re-emitted without energy loss but with a possible change in direction [8]. Although the changes in direction can be small for single scattering events, multiple scattering events can lead to a randomisation of the photon propagation direction [8]. This in turn causes image blur, thereby reducing image resolution. To quantify the influence of absorption and scattering on the propagation of photons through tissue, a physical parameter called the mean free path (MFP) is used. The MFP quantifies the average distance a photon travels between two consecutive scattering events, and it can be expressed as:



$$MFP = \frac{1}{\mu_s + \mu_a} \quad (5.1)$$

where  $\mu_s$  and  $\mu_a$  are the scattering and absorption coefficients, respectively, which describe the probability of photon scattering and absorption in tissue per unit path length. For biological tissue,  $\mu_s \gg \mu_a$ , therefore the MFP can be approximated by:

$$MFP = \frac{1}{\mu_s} \quad (5.2)$$

The MFP and its analog for describing multiple scattering events, the transport MFP (TMFP) are useful standards to compare the tissue penetration capabilities of different optical imaging techniques. The MFP not only depends on the optical properties of tissue, i.e. highly opaque tissue such as muscle or skin will scatter strongly, whereas semitransparent tissue such as the skin of zebrafish will scatter weakly, but also on the wavelength used for imaging. Longer wavelengths (the near infrared) generally scatter less than shorter wavelengths (UV and the visible spectrum). A detailed description of the MFP for different imaging techniques can be found in Ntziachristos' review [8].

### Fluorescence microscopy techniques

Despite being limited in their penetration depth into tissue, 3D light microscopy techniques address some critical shortcomings of medical tomographic methods described in the previous section by imaging samples at a smaller scale (tens of micrometres), with higher resolution (hundreds of nanometres) in 3D, by making use of sectioning to differentiate between objects separated along the optical axis. These techniques achieve sectioning either by physical means, such as using a microtome to cut thin tissue slices that can later be examined with a microscope [9]; or by optical means, which involve the rejection of any out-of-focus signal to acquire high contrast images of specific focal planes inside a thick sample [10]. In both cases, these individual slices are computationally combined into a 3D reconstruction of the sample.

An example of imaging using physical sectioning is High Resolution Episcopic Microscopy (HREM) [11]. In HREM, a thick tissue sample, up to  $5 \times 5 \times 7 \text{ mm}^3$ , is first embedded in a solid resin block. Images of the block face are captured with a widefield microscope as the block is sectioned by a microtome. The images of each slice are then stacked together to generate a 3D reconstruction of the sample [12]. The resultant images are inherently aligned and therefore require no digital transformations in post-processing, as they are acquired from the block face. The depth resolution in HREM is determined by the number of sections, which can be as thin as  $1 \text{ }\mu\text{m}$  [13], and the lateral resolution is determined by the NA of the

objective lens used to capture the images, with a value as low as 1  $\mu\text{m}$  previously reported [13]. HREM requires histological-type sample preparation and is therefore destructive and cannot be implemented with live samples. Additionally, the instrumentation required is expensive and not widely available (see Table 5.1).

Imaging using optical sectioning is far more common, and the most widespread and frequently used technique is Laser Scanning Confocal Microscopy (LSCM). LSCM enables the high resolution imaging of fluorescently labelled specimens by rejecting out-of-focus signal from the region of interest using a pinhole in a plane conjugate to the image plane [14]. The image quality and optical sectioning resulting from LSCM depends directly on the size of the pinhole used. The smaller the pinhole, the better the optical sectioning; however, this leads to a lower total signal collected which either degrades the signal-to-noise ratio, or results in long per-pixel dwell times. The larger the pinhole, the higher the resulting image contrast but both the lateral resolution and the optical sectioning deteriorate. Furthermore, since the acquisition process involves raster-scanning a laser spot across the sample at different focal planes, LSCM is slow, prone to causing photobleaching (elimination of a dye's ability to emit photons), and is typically limited to thicknesses of up to 100 micrometres [15].

To address the speed and photobleaching issues in LSCM, Selective Plane Illumination Microscopy (SPIM) [16] is an alternative fluorescence microscopy technique with optical sectioning. SPIM can generate high resolution 3D data from biological specimens up to a few millimetres in size, non-invasively. Optical sectioning is achieved by illuminating the sample with a thin sheet of light generated by a cylindrical lens or by a set of dithering mirrors [17], such that only fluorescence from this thin section is recorded on the detector. 3D images are then generated by scanning the sample laterally through the light sheet and imaging the fluorescent emission from the excited region in an orthogonal direction to the propagation of the illumination. The lateral resolution in SPIM is limited by the NA of the imaging objective, and the axial resolution is determined by the thickness of the light sheet. When using Gaussian beams to generate the light sheet, diffraction determines how thin the sheet can be, and as such, there is a trade-off between light sheet thickness and sample size. A larger sample requires either a thicker light sheet, which reduces the axial resolution, or additional scanning through the sample, which increases acquisition times. This trade-off has been addressed by using swept Bessel beams in the lattice light sheet design [18] or by the elegant field synthesis theorem to generate scanned light sheets [19], although both methods have been optimized to image groups of cells at high-resolution rather than large tissues. Moreover, both methods require beam-shaping elements, such as a spatial light modulator (SLM) to generate their light sheets, adding complexity and cost to their design. The penetration depth of SPIM is fundamentally limited by scattering, which

can be partially overcome by introducing additional excitation and detection optics [20]. This also increases the complexity and cost of implementation for imaging large and thick samples. An open-source implementation of SPIM, the mesoSPIM, uses tunable lenses to scan the thinnest portion of the light sheet across the sample and stitches the resulting images, allowing centimetre-scale samples to be scanned with near isotropic resolution [21]. The mesoSPIM is expensive (> £100,000), requires technical expertise to build and operate, and only 10 systems exist worldwide (<https://mesospim.org/>). Both LSCM and SPIM work by exciting fluorescent molecules with a focused laser beam; therefore, neither technique can detect signals from unlabelled tissue surrounding the fluorescent area.

### Hybrid imaging techniques

Two alternative optical imaging techniques which employ hybrid modalities to achieve much larger penetration depths, from a few millimetres up to centimetres, are Optical Coherence Tomography (OCT) and Photoacoustic Imaging (PAI).

OCT is a hybrid between light microscopy and tomographic techniques which allows non-invasive cross-sectional imaging of in-vivo biological samples by using low-coherence interferometry with gentle illumination in the near infrared [22]. In OCT, a fibre optic Michelson interferometer is used to illuminate the sample using near-infrared light (~1300 nm) and collect the reflected light from tissue, which is then combined with the reflection from the reference mirror. An interference signal is only detected when the group delay of the sample and the reference arm are almost matched (close to equal time-of-flight), and this signal falls off steeply with the delay mismatch due to the low-coherence nature of the light source. This allows the longitudinal location of the reflected tissue signal to be detected with high resolution. Volumetric images can then be built up after collecting neighbouring depth scans of the sample. OCT allows for lateral resolutions of 10-20  $\mu\text{m}$  and axial resolution of ~7  $\mu\text{m}$  [23], however it only permits a limited penetration depth (1-3 mm) [23] due to strong absorption and scattering events beyond this length. This technique is routinely used for retinal imaging in both clinical and research settings. Its limited penetration depth makes it unsuitable for whole organism imaging except in the very early stages of embryonic development [24].

PAI is a hybrid between optical and ultrasound imaging in which tissue is irradiated with short (nanosecond) laser pulses either in the visible or the NIR, leading to absorption by chromophores (e.g. water, lipids, haemoglobin, etc.) or exogenous contrast agents [25]. The absorbed radiation is rapidly dissipated in the form of heat, which creates a pressure differential that results in the emission of broadband, low-amplitude acoustic waves. These acoustic waves propagate to the sample surface, where they can be detected using

ultrasound transducers and can be used to reconstruct a 3D map of the absorbed optical energy distribution[26]. Unlike in traditional ultrasound, in which image contrast depends on the mechanical properties of tissue, in PAI contrast depends on the optical absorption properties of tissue. Since differences in optical absorption between different tissues are often much larger than those in mechanical properties, PAI offers greater specificity than traditional ultrasound to differentiate different types of tissue [27]. PAI can achieve imaging depths of several centimetres when operating in the NIR window (620-950 nm), with resolution of a few hundred micrometres, although higher resolution can be achieved by reducing the imaging depth. This technique has been primarily applied in cancer research, as it is useful for tumour detection [28] and assessment of therapeutics against cancer progression [29], among other uses.

### **Applicability to living systems**

The techniques introduced above have varying compatibilities with live organism imaging. Medical tomography techniques such as CT and MRI are routinely used to image whole live animal and human bodies. Although exogenous contrast agents can be used to improve image contrast, these are not required, and both techniques can image dense tissue non-invasively. Fluorescence microscopy techniques can image live cells and small, semi-transparent organisms such as fruit fly embryos, zebrafish embryos, or *C. elegans*, however for imaging thicker samples ( $> 100 \mu\text{m}$ ) such as brain or lung tissue, optical clearing is required to reduce the influence of scattering on image quality. Optical clearing is a chemical process in which organic solvents are used to make biological tissue transparent by matching the refractive index of the tissue layers to that of the solvent [30, 31, 32]. For clearing, tissue is first fixed to hold the proteins of interest in place, and delipidised to reduce scattering. The fixed tissue is then dehydrated, as the refractive index of water is significantly lower than that of cellular structures with proteins and lipids. Following dehydration, the tissue is immersed in an organic solvent, such as a solution containing benzyl alcohol and benzyl benzoate (BABB), which chemically matches the refractive index of the remaining tissue structures with the solvent. This results in the sample being effectively transparent, such that incoming photons can excite fluorescent emitters deep within the sample, and the emitted fluorescence can reach the detector [32]. Clearing is a highly destructive process; therefore, it cannot be performed with live samples. For more information on different tissue clearing protocols and the underlying chemistry of the process, Tainaka et al. [33] offer a comprehensive review. OCT and PAI, on the other hand, use laser excitation in the NIR, where light toxicity is low and penetration depth is highest for biological tissue, and can therefore be used in living samples. OCT is routinely used to image cross-sectional slices of retinas in live animal and

human eyes [23]. PAI can achieve even higher penetration depths in dense tissue than OCT, as the acoustic signal collected scatters ~1000 times less than visible light and can therefore propagate longer distances in tissue with minimal attenuation [27]. PAI has been used to image tissue both in live small organisms [29, 34, 35, 36, 37], and in humans [27, 38].

In summary, the medical tomography techniques (micro-CT and micro-MRI) introduced above can visualise morphology in large samples (tens of centimetres) with micrometre-level resolution and are therefore useful for whole-body imaging, however they are prohibitively expensive for most research labs and cannot match the high resolution and specificity available to fluorescence microscopy. 3D fluorescence microscopy techniques (LSCM, SPIM) can visualise features at sub-micrometre resolution, but they cannot image samples thicker than a few hundred micrometres due to tissue scattering. The penetration depth of fluorescence microscopy can be extended by optically clearing the sample, which can only be done with inert, fixed tissue. HREM can handle larger specimens than LCSM and SPIM (5x5x5 mm<sup>3</sup>) with similar resolution, however its sample preparation protocol is complicated, invasive, and destructive. OCT can also achieve micrometre-level isotropic resolution, with penetration depths in the range of 1-3 mm, whereas PAI can achieve penetration depths of centimetres, at a resolution of a few hundred micrometres. Table 5.1 below shows a summarising comparison of the techniques introduced above.

### 5.1.3 Optical Projection Tomography

A bridging technique between the tomographic and microscopy techniques presented above which can both examine morphological features and use fluorescence labelling in a sample is Optical Projection Tomography (OPT). OPT was first proposed by Sharpe et al. in 2002 [42]; it is a 3D imaging technique designed for transparent mesoscopic samples which enables the visualization of micrometre-scale features. OPT addresses the sample scale gap between the tomographic techniques (samples > 20-30 mm), and light microscopy techniques (samples < 1 mm), to image biological samples in the 1-20 mm regime.

OPT is based on computerized tomography techniques [43] in which 2D images, called projections, are acquired at different sample orientations, and then used to obtain a 3D volume of the sample using a reconstruction algorithm, such as filtered-back projection (FBP). Sample clearing [33] is often necessary to allow light propagation and imaging through the thickness of the sample. OPT can operate using either fluorescence (emission OPT, eOPT [42]), or absorption/scattering of the sample (transmission OPT, tOPT [42]) to generate image contrast. In eOPT, a fluorescent sample is excited using a wide-field illumination source. The photons emitted by a fluorophore of interest are collected by a detector with an appropriate filter to reject the incident illumination. In tOPT, a source directs

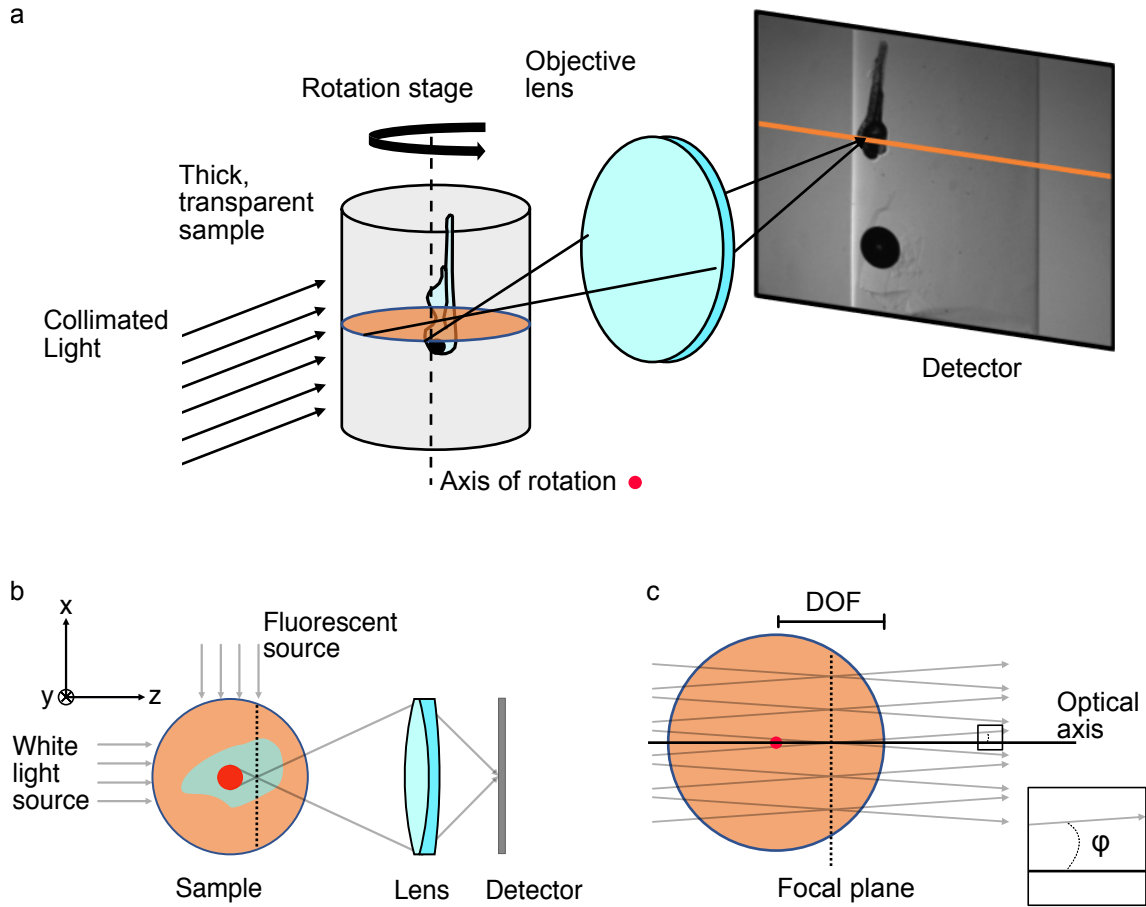
**Table 5.1** Comparison of volumetric imaging techniques.

Method	Sample size (width · height · depth)	Resolution	Live	Cost (£)	Comments & Source
micro-CT	140 x 88 x 140 mm <sup>3*</sup>	~45 µm isotropic	Yes	> 100K	Bruker SkyScan2214
micro-MRI	160 x 72 x 160 mm <sup>3*</sup>	~60-120 µm (isotropic)	Yes	> 100K	Bruker PharmaScan 7T
HREM	5 x 5 x 7 mm <sup>3</sup>	~1-5 µm (isotropic)	No	> 100K	Chapter 3, Geyer et al. 2017 [13]
LSCM	160 x 110 x 0.28 mm <sup>3</sup> (with a scanning stage), 135 x 135 x 280 µm <sup>3</sup> (still), depth up to 5 mm for cleared sample	~240 µm (lateral), ~700 nm (axial)	Yes	> 100K	Zeiss LSM980
SPIM (home built)	1 x 1 x 1 mm <sup>3</sup>	~1 µm (lateral), ~6 µm (axial)	Yes	< 100K	Power and Huiskens 2017 [39]
SPIM (commercial)	10 x 10 x 20 mm <sup>3</sup> (cleared sample)	~1.8 µm (lateral), ~7 µm (axial)	Yes	> 100K	Zeiss Lightsheet 7
OCT	12 x 9 x 3 mm <sup>3</sup>	~10-20 µm (lateral), ~7 µm (axial)	Yes	> 50K	Chapter 3, Aumann et al. 2019 [23]
PAI	15 x 15 x 3 cm <sup>3</sup> (large FOV, commercial), 1 x 1 x 3 mm <sup>3</sup> (small FOV)	large FOV: ~150 µm (isotropic), small FOV: ~45-120 µm (lateral), ~15 µm (axial)	Yes	~470K	Binte et al. 2019 [27], iThera MSOT inVision 128
macro-OPT	60 x 45 x 45 mm <sup>3</sup> (cleared sample)	~50 µm <sup>**</sup>	No	< 5K	Lee et al. 2017 [40]
OptiJ (this thesis)	8 x 8 x 13 mm <sup>3</sup> (cleared sample)	~50 µm <sup>**</sup>	No	< 5K	Vallejo Ramirez et al. 2019 [41]

\* Max sample diameter is determined by bore size; in Section 5.3.2.

\*\* Resolution in OPT is spatially variant, and is discussed

quasi-collimated, uniform illumination onto the sample, as shown in Fig. 5.1. The intensity collected at the camera is a function of the amount of light absorbed by the sample. In other words, the image recorded by the detector is a projection of the attenuated radiation that traverses the sample. These two modalities can work together to display fluorescent signals in the context of the whole sample. The comparison between two OPT implementations, macro-OPT [40] and OptiJ [41], the latter of which is presented in Chapter 6 of this thesis, is shown in Table 5.1.



**Fig. 5.1** (a) Schematic of a general OPT setup. The sample is embedded in a clear, solid medium and only rotated, not translated, as sequential projections are acquired on a fixed detector. Collimated light which is transmitted through the sample is collected by an objective lens and focused onto a detector. (b) Top view of the schematic in (a) showing the illumination path for eOPT with a fluorescent source, and tOPT with a white light source. (c) The focal plane of the objective lens is placed halfway between the axis of rotation and the front of the sample. The NA of the lens, represented here by the half angle of the imaging cone  $\phi$ , is small such that the depth of field is sufficiently large to span the front half of the sample. This ensures at least one half of the sample will always be in focus at the detector.

## 5.2 Principles of Optical Projection Tomography

This section provides an explanation of the working principles behind OPT and a brief overview of the mathematical basis of tomographic reconstructions. Artefacts resulting from the reconstruction process that affect image quality are discussed, as well as strategies with which to mitigate their effects. There will also be an overview of studies carried out by different research groups on optimizing the acquisition speed, reconstruction resolution, and reconstruction speed for OPT, along with some interesting biological applications.

### 5.2.1 Overview of OPT working principle

OPT uses a finite conjugate imaging system with an objective lens and a detector – similar to a microscope— which means detected ray bundles have a finite angular spread as shown in Fig. 5.1 b and c. This angular spread defines the depth of field (DOF), which is the acceptable movement range of the detector or the object along the optical axis such that a small enough blur is produced in the recorded image. For objectives requiring a large DOF, the angular spread of a ray bundle will be small. Therefore, the light rays collected at the detector will approximate parallel lines through the sample. This approximation only works when the sample volume is within the DOF of the imaging system. The DOF is labelled in Fig. 5.1 c for a set of ray bundles traversing a horizontal cross-section of the sample. There is an important trade-off between imaging resolution and DOF, i.e. between the smallest features which can be detected and the thickness of the sample that can be imaged. This trade off can be seen directly from the equations relating the NA of the objective lens to the resolution of the system by the Abbe limit [44] for incoherent imaging:

$$\Delta x = \frac{\lambda}{2NA} \quad (5.3)$$

and to the DOF:

$$DOF = \frac{2\lambda}{NA^2} \quad (5.4)$$

where  $\Delta x$  is the transverse size of the smallest feature that can be resolved in the sample and  $\lambda$  is the wavelength of the light. As the NA increases ( $\Delta x$  becomes smaller), the DOF decreases, and subsequently the point spread function (PSF) invariance decreases along the optical axis. It is common practice in OPT systems to balance this trade-off by locating the focal plane a quarter of the way into the sample such that the DOF covers the front half of the sample [45], as shown in Fig. 5.1 c. In this way, the whole extent of the sample will be



imaged in focus through a full rotation. This also means every projection will contain an out-of-focus contribution from the second half of the sample.

In tOPT, rays originating from the illumination source are collimated by a lens and travel through the sample, scattering or attenuating as they traverse the volume and reach the detector. This modality is similar to conventional x-ray CT, in which non-diffracting beams travel through the sample in straight lines to reach the detector. For OPT, light rays do indeed diffract, however the sample features intended to be resolved are much larger than the wavelength of the light (tens of micrometres) and therefore diffraction is not dominant. In eOPT, the light sources are the fluorescent emission regions inside the sample that emit incoherent photons in all directions, some of which are focused onto the detector. The image formation process in OPT is incoherent, either with light transmission in tOPT or with fluorescence emission in eOPT. In addition to having ray bundles with a small angular extent by limiting the NA of the objective lens, a telecentric lens can be used in both tOPT and eOPT to capture photons travelling approximately parallel to the axis, such that their path approximates straight lines. This approximation allows the use of the same equations for reconstruction of traditional x-ray tomography images to be used in both tOPT and eOPT. The following section will describe the most commonly used reconstruction method, filtered back-projection (FBP), as it is also the method used to reconstruct the tomographic data presented in Chapter 6. Other reconstruction techniques using iterative methods, convolutional neural networks, or deconvolution will be briefly mentioned in section 5.3.

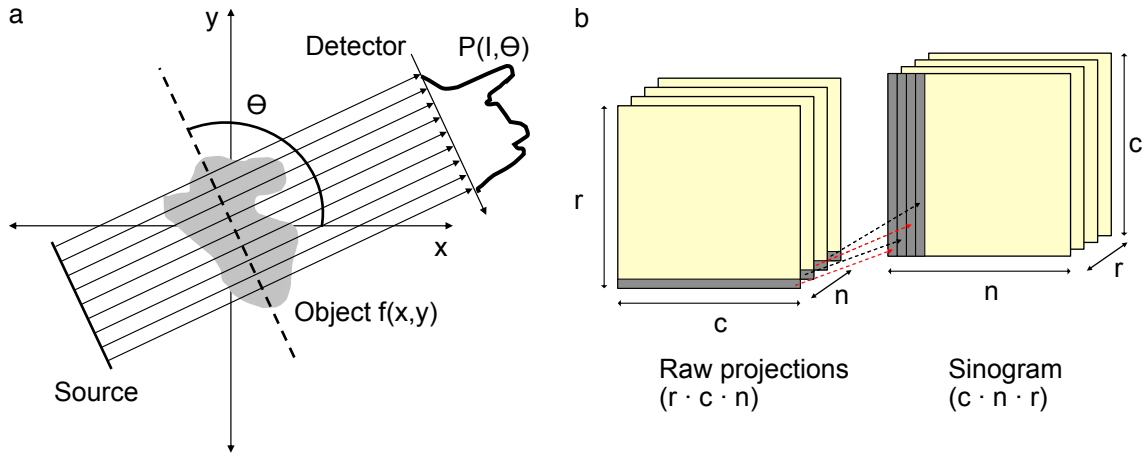
### Reconstruction via filtered back-projection

The reconstruction of the cross-sectional slices of an object from its projections is analogous to solving the inverse problem of determining the shape of an object using its shadows when illuminated at different angles. In the case of OPT, orthographic projections are used to reconstruct cross-sectional slices of a sample, which can be stacked up to recover the 3D volume of the sample. FBP is the most common algorithm used to achieve this reconstruction. This section describes the principles behind FBP and its implementation.

In OPT, if the axis of rotation of the sample is parallel to the imaging plane of the camera, then a single horizontal slice from the object will map directly onto a single row of pixels for every angle through a full rotation. Projections are then intensity profiles for a given angle of rotation  $\theta$ , essentially a collection of line integrals for a constant  $\theta$ . These line integrals represent the total attenuation experienced by an illumination ray as it travels through the object. Mathematically, a projection for an object  $f(x, y)$  can be expressed as

$$P(I, \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(x \cos \theta + y \sin \theta - I) dx dy \quad (5.5)$$

where  $P(I, \theta)$  is known as the Radon transform of the function  $f(x, y)$  [43], as illustrated in Fig. 5.2 a, and  $\delta(x \cos \theta + y \sin \theta - I)$  defines a set of parallel lines (representing rays) normal to the detection plane. The plot of the Radon transform data is known as a sinogram, since the Radon transform of an off-center point source is a sinusoid. To create a sinogram, every pixel row corresponding to the same horizontal 2D plane in the sample is grouped vertically as a function of angular step, as shown in Fig. 5.2 b. This results in a plot of  $n$  projection slices with  $c$  pixels stacked adjacent along the horizontal axis, where  $c$  is the number of original pixel columns in the detected image.



**Fig. 5.2** Schematic illustrating the Radon transform operation. (a) Orthographic projections of an object are obtained at the Detector by collecting parallel rays that are attenuated as they traverse through the object. (b) Schematic of the permutation operation to create a sinogram. From the stack of  $n$  raw projections with dimensions  $r \times c$  (corresponding to the height and width of the detector in pixels), every row in the stack that corresponds to the respective horizontal plane in the sample (horizontal dark grey bars) is grouped vertically as a function of angular step. This results in a plot of  $n$  projection slices with  $c$  pixel rows stacked one next to the other along the horizontal axis (dark grey bars). The output is then a stack of  $r$  images with dimensions  $c \times n$ , in which each slice represents a sinogram corresponding to one horizontal 2D plane in the sample.

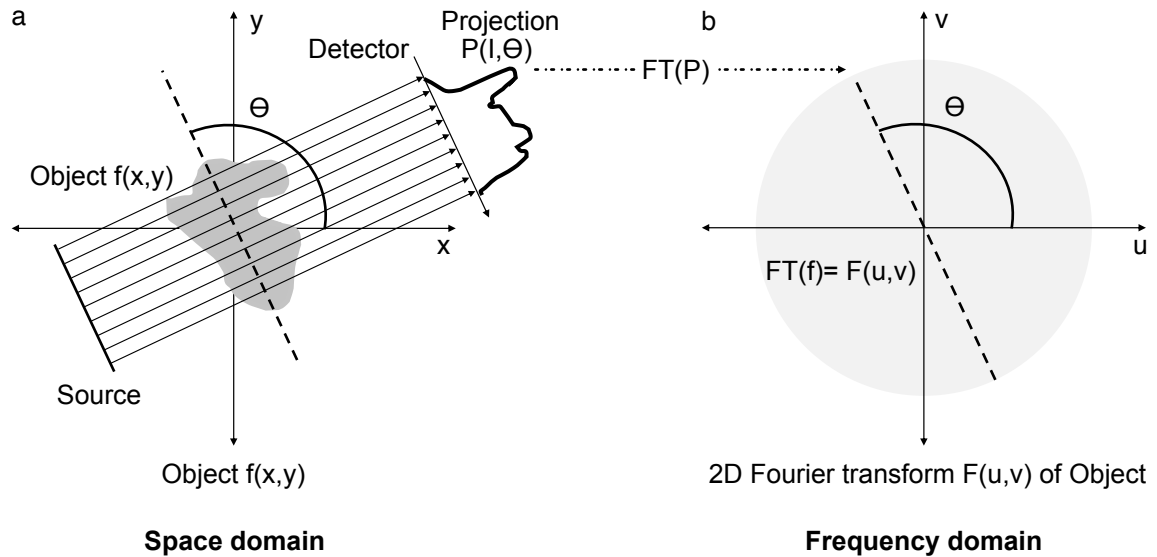
In practice, a finite set of  $n$  projections are acquired over a full rotation with an angular step given by

$$\Delta\theta_n = \frac{360^\circ}{n} \quad (5.6)$$

such that a given row in a projection contains information from a single horizontal 2D plane at the corresponding angle of projection. The recovery of the cross-sectional slices of the object from its projections using FBP is possible due to the Fourier Slice Theorem.

### Fourier slice theorem as a basis for reconstructions

The Fourier Slice Theorem is a fundamental concept in the reconstruction of tomography data, as it relates the one-dimensional Fourier transform of a projection of an object function  $f(x,y)$ , to the two-dimensional Fourier transform of  $f(x,y)$  which we will define here as  $F(u,v)$ , where  $u,v$  are the frequency domain variables corresponding to the space domain variables  $x,y$ . The Theorem states that the Fourier transform of a projection, which is the line integral of the attenuation profile at an angle  $\theta$ , is equal to a slice through the 2D Fourier transform of the original object at the same angle  $\theta$  through the origin (from Kak and Slaney, Ch. 3 p.56 [43]). Therefore, by collecting a set of projections  $P(I, \theta)$ , their individual 1-D Fourier transforms can be assembled together to fill the  $uv$ -plane (frequency domain) with radial slices that can then be inverse Fourier transformed to recover the original shape of the object. The Fourier Slice theorem is illustrated in Fig. 5.3. Kak and Slaney, Ch.3, p.63 [43] provide a thorough explanation of the Fourier Slice Theorem.



**Fig. 5.3** Illustration of the Fourier Slice Theorem, which relates the 1D Fourier transform of a projection to a slice in the 2D Fourier transform of the original object. (a) Projection obtained at angle  $\theta$  from the Object  $f(x,y)$ , which is a line integral of the attenuation profile of the illumination rays through the sample. The 1D Fourier transform of the projection is equivalent to a radial slice at the same angle  $\theta$  through the 2D Fourier transform  $F(u,v)$  of the Object  $f(x,y)$ , as shown in (b), where  $u,v$  are the frequency domain variables corresponding to space domain variables  $x,y$ . That is, the Fourier transform of  $P(I, \theta)$  gives the values of  $F(u,v)$  along the dashed line in the  $u,v$  plane.

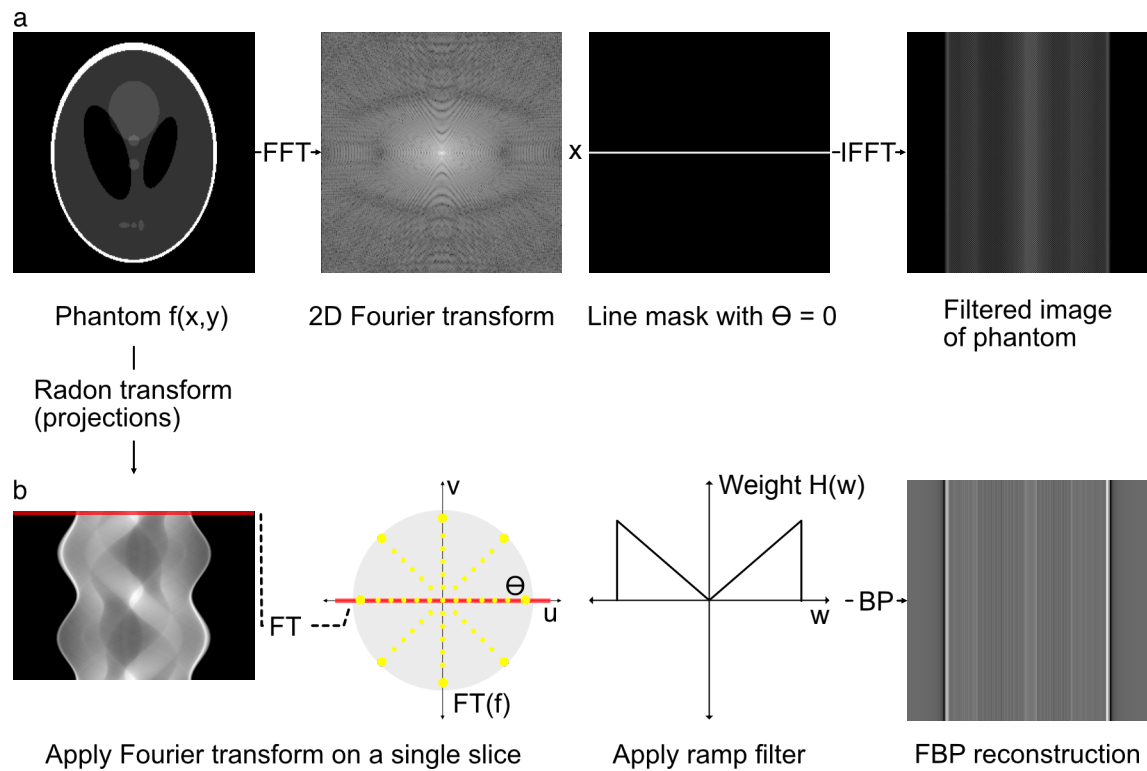
In other words, a 2D reconstruction of an object can be obtained by acquiring projections over a full rotation, taking the 1-D Fourier transform of each projection, using the Fourier

Slice Theorem to assemble them in the correct order in frequency space, and inverting the assembled points from the frequency domain back to the space domain. The 2D reconstructions obtained for each cross-sectional slice of an object can then be stacked up to recreate a 3D volume of the object. In practice, since only a finite number of projections are acquired, the assembly of 1D Fourier transforms of projections as radial slices into the  $uv$ -plane requires interpolation steps prior to performing an inverse Fourier transform to recover the object shape. This process is computationally intensive and may lead to unstable solutions, as stipulated by Rosenfeld and Kak, Ch. 8 p.367 [46]. Instead, most tomography applications implement a very accurate and fast algorithm for reconstruction: filtered back-projection.

### Filtered back-projection

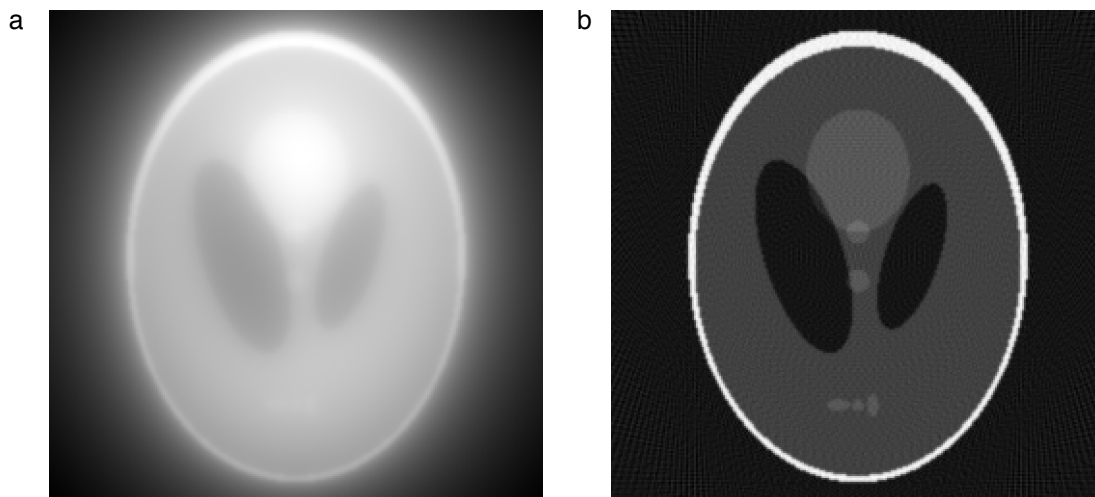
Before exploring the two main steps in FBP, a conceptual explanation can be put forth from the premise that each projection acquired represents an almost independent measurement of the object. This is apparent in the frequency domain following from the Fourier Slice Theorem, since the only common information in the Fourier transforms of two projections at different angles is the DC term [43]. The generation of a single projection can be thought of as a 2D filtering operation, since the Fourier Slice Theorem states that you can multiply by a line through frequency space to get a 2D projection, which is the equivalent to convolving by the FFT of a line. As a result, a distorted reconstruction can be produced by assuming all other projections are zero, positioning the values of the Fourier transform of the single projection in their respective place in the object's 2D transform, and performing an inverse Fourier transform. This procedure is equivalent to multiplying the original Fourier transform of the object by a line filter at the same angle with the coordinate axes as the projection ([43] Ch.3 pp.60-62), and it is illustrated in Fig. 5.4 a.

A simulation was performed in Matlab in which the Fourier transform of the 2D Shepp-Logan phantom [47], a standard test image for image reconstruction algorithms, was multiplied by a horizontal line filter in frequency space, and then inverse Fourier transformed to recover a distorted smear of the object (Fig. 5.4 a). The result of this operation was compared to a standard FBP reconstruction using the Matlab function *iradon* with only the first projection ( $\theta = 0$ ) of the Shepp-Logan phantom (Fig. 5.4 b). The results of these two operations in the right-most top and bottom row are very similar, save for the differences in contrast due to the discrete Fourier transform procedure for the top-row operation, and the extra filtering step performed during FBP in the bottom-row. This simulation demonstrates the conceptual explanation for FBP, and now we can consider the main steps involved in this reconstruction algorithm.



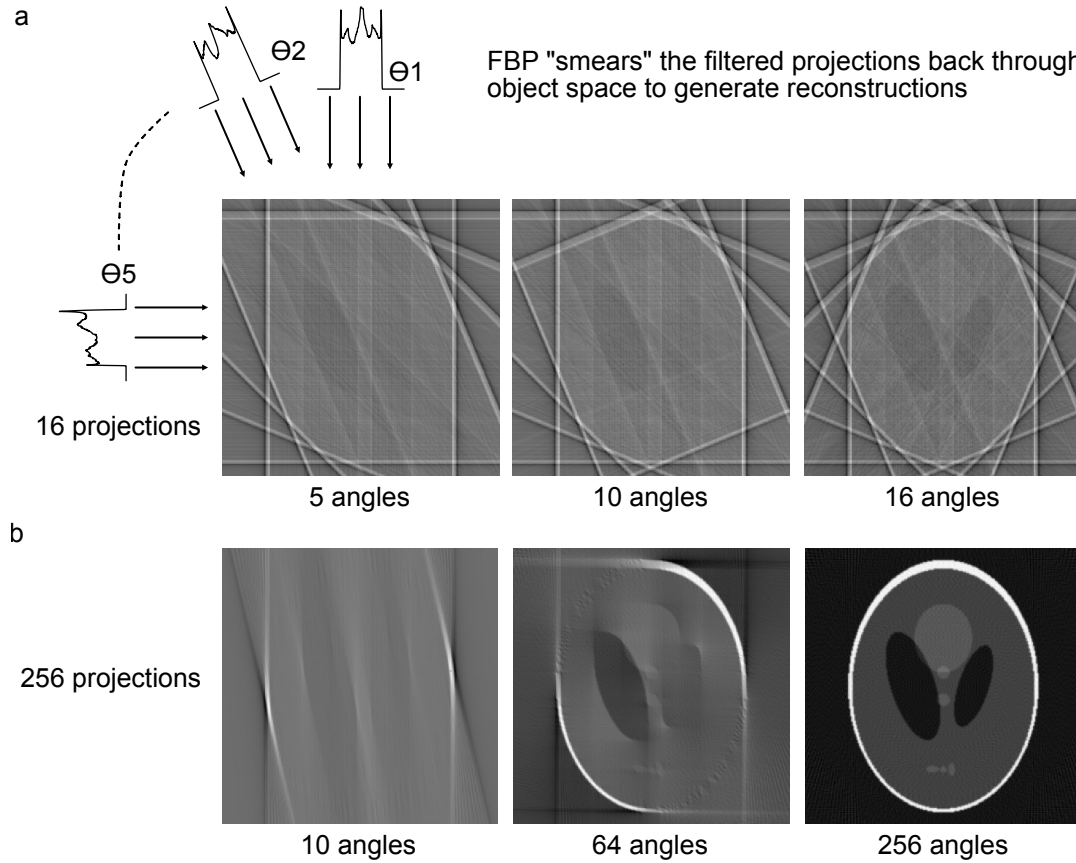
**Fig. 5.4** The measurement of a single projection can be understood as a 2D filtering operation. The Shepp-Logan phantom is used to demonstrate how both the FBP reconstruction from a single projection (a) is analogous to multiplying the original Fourier transform of the object by a line filter at the same angle with the coordinate axes as the projection (b). This simulation was performed using the script `part2Ch4_FBP_simulations.m`, which can be found in the code repository for this thesis <https://github.com/pedropabloVR/supplementary-code-thesis>.

FBP consists of two parts: a filtering operation, which involves a weighting of each projection in the frequency domain to avoid oversampling of the low spatial frequency information, and a back-projection part, which involves a summation of the inverse Fourier transforms of the weighted projections in the object space domain. The effect of the filtering step is illustrated in Fig. 5.5, in which a back-projection of the Shepp-Logan phantom is shown with and without filtering. Filters in both FBP for OPT and traditional x-ray CT seek to accomplish the same goal: to accentuate high-spatial frequency features and mitigate the effect of pixel "smearing" during back-projection. In the noiseless case, the ramp filter (Ram-Lak [48]) ideally compensates for the oversampling of the slices near the axis of rotation and yields a high-quality reconstruction. In practice, real tomography data has noise, and thereby any high-pass filter will also accentuate the noise components in the reconstruction. The ramp filter is sensitive to noise; therefore, it can be multiplied by a sinc function (Shepp-Logan filter) or a Hamming window (Hamming filter) to reduce accentuation of high-frequency noise. In principle, OPT and CT can use the same filters for FBP. In practice, however, modern CT systems use iterative reconstruction (IR) methods instead of FBP, which either iteratively try different filters to reduce noise in image space after back-projection, or iteratively cycle between forward and back projections until a continuously-updated filter attenuates the noise under a given threshold [49]. OPT systems still largely use FBP and common filters (Shepp-Logan, Hamming, etc.), however new reconstruction methods which reduce noise have recently been reported and will be discussed in Section 5.3.



**Fig. 5.5** Reconstruction of the Shepp-Logan phantom with back-projection over 256 angles without (a) and with (b) a ramp filter.

The latter step is known as back-projection because it can be thought of as “smearing” each of the acquired projections back through the image plane to reconstruct the image of the object, as shown in the first image panel of Fig. 5.6 a.



**Fig. 5.6** OPT reconstruction via FBP can be visualised as smearing the filtered projections back through the image plane. (a) Image panels show different junctures in the FBP reconstruction of a Shepp-Logan phantom using 16 projections. Junctures for 5, 10, and the full 16 angles are shown. The first panel illustrates how the projection data for each  $\theta_n$  orientation are sequentially added back onto the image plane in order to recover a faithful image of the original object, similar to smearing the projection values across the image plane. (b) Reconstruction using 256 projections with junctures at 10, 64, and the full 256 angles to demonstrate how a larger number of projections leads to better quality images. This simulation was performed using the script `part2Ch4_FBP_simulations.m`, which can be found in the code repository for this thesis <https://github.com/pedropabloVR/supplementary-code-thesis>.

Mathematically, the filtering and back-projection steps can be expressed as:

1. Filtering in frequency space by multiplying the Fourier transform of each projection by an appropriate filter (e.g. Ramp, Shepp-Logan, or Hamming):

$$f_{filt}(x,y) = FT\{P(I, \theta_n)\} \cdot H(w) \quad (5.7)$$

where  $P(I, \theta_n)$  is a projection for angle  $\theta_n$  and  $H(w)$  is a ramp filter in frequency space, with  $w$  having units of spatial frequency.  $w$  is related to the  $u, v$  plane (Fig. 5.3 and 5.4) by  $u = w \cos \theta$  and  $v = w \sin \theta$  (Kak and Slaney, Ch.3, p.63 [43]).

2. Performing back-projection over the full rotation of the sample:

$$R_{fbp}(x,y) = \int_0^{2\pi} f'_{filt}(x \cos \theta + y \sin \theta, \theta) d\theta \quad (5.8)$$

where  $f'$  is the inverse Fourier transformed filtered projection data and  $R_{fbp}$  is the back-projected image (Kak and Slaney, Ch.3, p.63 [43]).

Since the data are only sampled at intervals of  $n$ , which leads to discrete sampling of the Fourier space representation of the object by the Fourier Slice Theorem, an additional interpolation step is needed before back-projection [43]. A complete explanation of FBP can be found in Chapter 3 of Kak and Slaney [43], along with rigorous derivations and proofs of the mathematics of tomographic reconstruction. Now that a description for the reconstruction procedure for tomographic data has been introduced, the practical shortcomings and resulting artefacts from this reconstruction procedure will be discussed.

### Artefacts during reconstructions

OPT reconstructions often suffer from artefacts which degrade image quality and distort the resulting cross-sectional slices of the sample. The most common reconstructions artefacts and their cause will be briefly discussed here, along with strategies to mitigate their effects.

For OPT imaging of large mesoscopic samples, in the range of tens of millimetres, uniform illumination can be a challenge. Uneven illumination, due to an intensity fall-off at the edges of the field from the collimated light source, can lead to shadowing and dark corners in the tOPT projections. The Beer-Lambert Law [50] can be used to correct for uneven illumination artefacts; each projection is normalised by a low-noise estimate of the illumination profile obtained by averaging a set of calibration brightfield images acquired without a sample [41]. An example of this correction will be shown in the following chapter, as part of the open-source software developed. For eOPT, uneven illumination results in reduced fluorescence excitation (and therefore less signal being collected) away from the

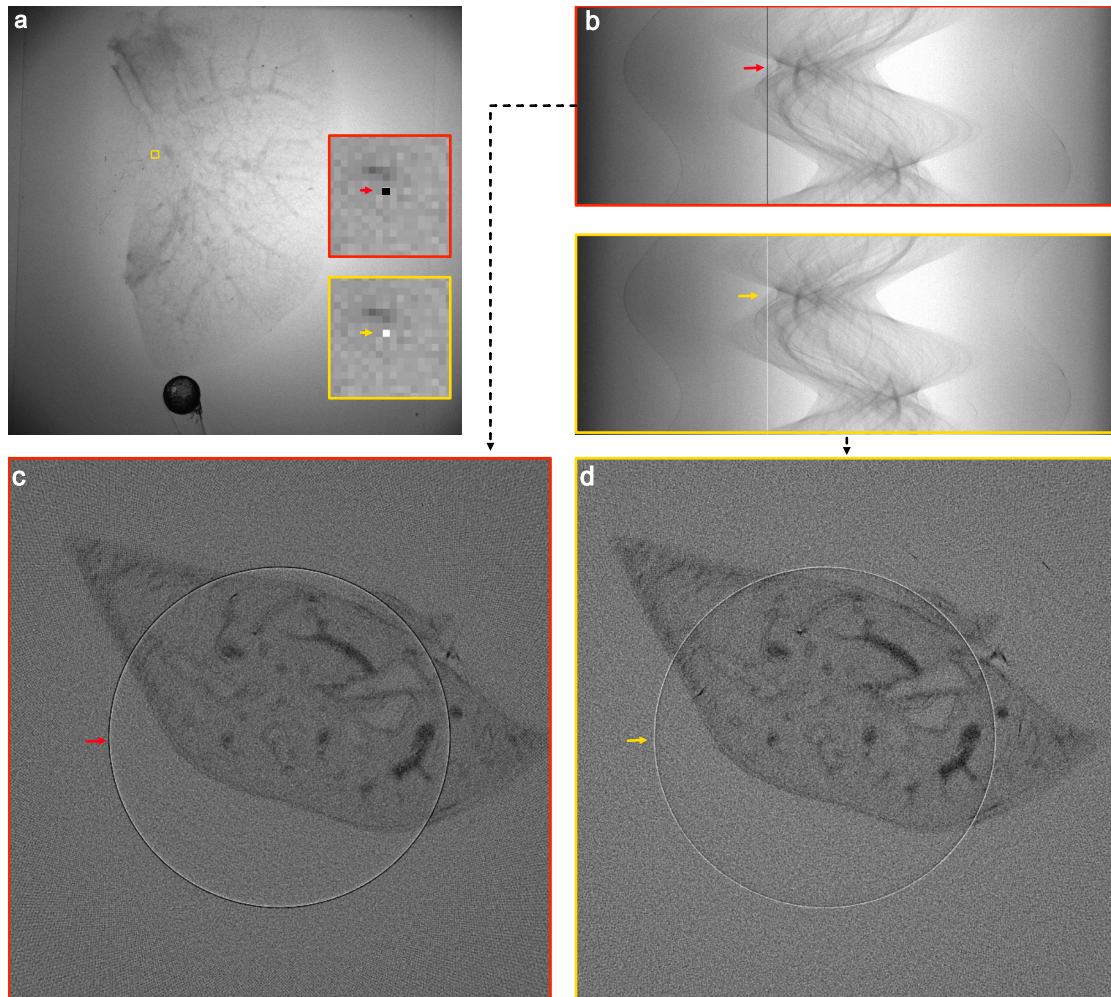


illumination's optical axis, such that features in off-axis positions are dimmer than features close to the axis. This can be addressed using oversized optics to collimate the light source such that the dark corners fall outside the field of view [41], at the cost of reduced throughput.

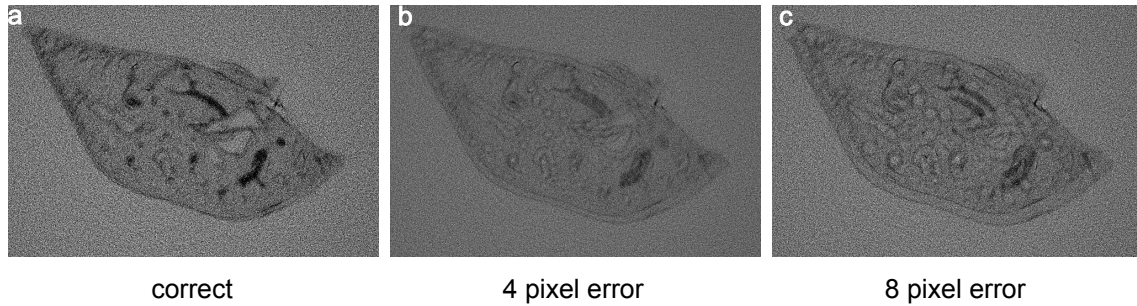
Due to the radial nature of the acquisition process, correlated noise within the detector can also cause artefacts in the reconstructions. Sensor blemishes such as “hot” pixels, which appear completely white (pixels with a high dark current, often in an uncooled detector), and “dead” pixels, which appear completely dark, can cause concentric ring artefacts in OPT reconstructions, such as those shown in Fig. 5.7 c and d. These blemishes can also be easily detected in the sinogram for the projections, as they appear as solid black or white lines as shown in Fig. 5.7 b. These artefacts can be corrected for by replacing the intensity of the bad pixels by the mean of its 8 neighbours, or another similar interpolation procedure [51]. Alternatively, a cooled sensor may be used to mitigate the occurrence of these sensor blemishes.

An incorrect estimation of the axis of rotation during reconstruction can lead to halo-like artefacts in the reconstructed cross-sectional slices (Fig. 5.8 b and c). If the position of the axis of rotation of the sample is not well calibrated prior to acquisition, then it must be estimated during the reconstruction to mitigate the effects of its misalignment. Users can make an educated guess of the position of the rotational axis during reconstruction of a sinogram and refine their guess by choosing the highest quality image, subsequently reconstructing the sinogram many times until the best reconstruction is found empirically. This iterative process is time-consuming, and becomes infeasible for large data sets, therefore two computational approaches are available: Walls et al. [51] use an autofocus-style measure in which a set of images is reconstructed with a range of positions for the axis of rotation, and the global variance in pixel grey-levels is calculated for each reconstruction. The resulting image with the highest variance (least blurring) is the closest to the ideal reconstruction. Vallejo-Ramirez et al. [41] fit the trajectory of a fiducial glass bead embedded close to the sample to determine the position of the axis of rotation, and to measure how misaligned it is from the ideal position perpendicular to the optical axis and parallel to the plane of the camera. The advantage of Walls et al.'s image variance method [51] is that it does not require a fiducial marker in the sample, which makes sample preparation easier and allows more of the field of view to be used for the sample, however reconstructions are more time-consuming, as each slice is reconstructed multiple times. The glass bead method adds a step to the sample preparation method, however the estimation of the axis of rotation position can be used for the entire sample so it does not lengthen reconstruction times.

This subsection described the most prevalent artefacts found in OPT reconstructions. A more comprehensive explanation of artefacts in OPT reconstructions can be found in



**Fig. 5.7** Hot or dead pixels on the sensor cause ring artefacts in the reconstructions. (a) A projection of a murine lung with an artificially induced dead (red arrowhead in red box) and hot (yellow arrowhead in yellow box) pixel to simulate the effect of detector imperfections on the reconstructions. (b) The sinogram corresponding to the horizontal row of pixels containing the “bad” pixel will exhibit a solid black (red arrow) or solid white (yellow arrow) line through it, as the location of the pixel does not vary as the sample rotates. When the sinogram undergoes FBP, this “bad” pixel results in a ring artefact in the sample, almost identical for the case of the dead (c) and the hot (d) pixel. The images were reconstructed in ImageJ using the OptiJ plugins [41].



**Fig. 5.8** An incorrect estimation of the sample's axis of rotation during sinogram reconstruction leads to halo-like artefacts. A murine lung sample from the freely-available Vallejo-Ramirez et al. [41] data set is reconstructed with a correct estimate of its rotational axis (a), an incorrect estimation by 4 pixels (b) and 8 pixels (c) to demonstrate the degradation of image quality as the error increases.

Walls et al. [51]. These artefacts and other shortcomings of OPT, such as long acquisition times, small fields of view, and low resolution have been addressed by multiple groups [52, 53, 54, 55, 56, 57, 58, 59, 60] in an effort to scan larger volumes faster and with better resolution; some of these advances are described in the next section.

## 5.3 Advances in OPT

Since the first OPT publication by Sharpe et al. [42], there have been several efforts to apply OPT to many samples beyond mouse embryos, and to improve technological aspects of the technique in terms of resolution, acquisition speed, and image processing and reconstruction. A few of these efforts from other groups are briefly described in this section.

### 5.3.1 Acquisition time reduction and compressed sensing

The acquisition of a sufficient number of projections to produce a high-quality reconstruction in OPT, as well as the camera acquisition times required to capture the fluorescent signal specifically in eOPT can be time consuming and therefore inadequate for high-throughput imaging applications, or in-vivo samples. This is problematic, since one of the most exciting applications of OPT in its early days was its application in characterising morphological phenotypes in large knockout lines of mouse embryos for the International Mouse Phenotyping Consortium [61]. It is tempting to acquire fewer projections in order to reduce the acquisition time, however this introduces streaking artefacts during reconstruction via FBP, such as those present in Fig. 5.6 which only used 16 projections of the Shepp-Logan phantom for reconstruction.

One approach to minimize these streak artefacts involves using an experimentally determined modulation transfer function (MTF) mask for filtering during back-projection. This was reported to produce high-quality reconstructions with only 90 projections (a quarter of the acquisition time compared to traditional filtered back-projection) [53]. A different approach uses compressed sensing, a method to reconstruct a significantly under sampled signal, to reconstruct sparse OPT data sets without compromising image quality [54]. This implementation uses an iterative reconstruction method based on the well-documented TwiST algorithm [62] to solve the inverse problem of finding an object from a sparse number of projections. As few as 30 projections acquired in a few seconds were used to produce reconstructions with comparable quality to those of a much more densely sampled data set with 800 projections using FBP. Reducing acquisition times in OPT allows researchers to perform longitudinal studies on embryos during early development stages, decreases the probability of sample motion during acquisition for in-vivo experiments, and reduces the light dose to avoid phototoxicity and photobleaching of fluorescent proteins in the sample. This reduction in acquisition time allowed by compressed sensing comes at the cost of increased computation time, as iterative algorithms for compressed sensing could take over an hour compared to a few minutes for FBP with GPU acceleration [55].

### 5.3.2 Resolution improvement

In OPT, the resolution in the reconstructions depends on: the imaging resolution dictated by the NA of the objective lens used to acquire the projections; the pixel size which determines the thickness of the cross-sectional slices used to reconstruct the object; and the number of projections acquired, since an inadequate number of projections leads to image artefacts and reduces image quality. Furthermore, the residual invariance in the transfer function caused by the NA of the objective lens yields a spatially-variant post-reconstruction resolution, with components both in the radial direction from the rotational axis of the sample, and in the tangential direction from each radial position [57]. This is a result of the discrepancy between the straight-line geometry for projection acquisitions required by FBP, as is the case in X-ray CT, and the Gaussian-shaped ray bundle in OPT, which causes blurring on features separated from the focal plane. This blurring is manifested as a deterioration of the radial resolution with increased radius in the reconstructions.

Several approaches exist to improve the resolution in OPT, both during acquisition and post-processing. From the hardware perspective, to maximise the resolution during the acquisition of projections, a high NA lens can be used by limiting the imaging volume to less than a few millimetres, or an electrically tunable liquid lens can be used to scan the focal plane across the sample with a high NA while maintaining an effective depth of field covering

the first half of the sample with large volumes [52]. In post-processing, a deconvolution approach has been proposed which uses a spatially-varying PSF model as a function of the properties of the imaging optics which successfully recovers isotropic resolution in the reconstruction of a zebrafish [57]. This approach is advantageous, as the experimental PSF of the imaging system will vary significantly as a function of the field aberrations within the imaging volume. Whereas deconvolution with an invariant PSF is computationally simpler, this approximation can only be used when imaging small regions close to the optical axis. In the case of millimetre scale samples in OPT, the PSF will have contributions from both axial and transverse components which degrade the PSF shape away from the axis. A more recent approach uses a stretched Gaussian beam model to account for the discrepancy between ray geometry in x-ray CT and tOPT, and demonstrate higher contrast using their iterative reconstruction method compared to standard FBP for a non-sparse data set of a zebrafish head [58]. A novel method was also recently reported on the use of convolutional neural networks (CNNs) to produce streak-free reconstructions of under-sampled OPT data, faster and with fewer projections than in compressed sensing approaches. The CNN was trained with high-quality FBP reconstructions of densely sampled OPT data. This latter approach is promising for longitudinal and high-throughput 3D imaging of small organisms [55].

### 5.3.3 Applications of OPT

The use of OPT has been widely reported and its applications include the visualisation of 3D anatomy in mouse embryos [63, 64, 65, 56, 66, 67, 68, 69, 59, 24, 70, 60], zebrafish [71, 72, 73, 54, 74, 56, 75, 76, 60], fruit flies [77, 78, 79, 80], plants [81, 40], *C.elegans* [82], animal organs [83, 84, 67, 85, 59] and other mesoscopic samples [86, 87, 88]. In addition to undergoing improvements in resolution and acquisition speed, OPT has also been applied in combination with other imaging modalities such as SPIM [72, 59] to span a larger range of spatial scales that can be imaged, and FLIM [71, 89] to extract functional information from groups of fluorescent cells in a whole organism context. Despite its growing use, most OPT applications require advanced technical expertise to design and build the system [52, 40, 76, 90], expensive equipment [90], and bespoke software for reconstructions [90].

## 5.4 Summary

Large sample volumetric techniques such as micro-CT and micro-MRI are useful for exploring morphology in 20-160 mm sized samples such as organs, body parts, and small animals. Volumetric fluorescence microscopy techniques such as LSCM and SPIM can scan samples

from a few micrometres up to 2 millimetres, such as cell cultures and thin tissue slices with sub-cellular resolution, however they cannot image unlabelled structures and are limited in their penetration depth due to scattering.

OPT is a hybrid between the tomographic and fluorescence microscopy techniques discussed, as it can access information from both fluorescently labelled and unlabelled structures in sample, such that it can explore molecular-specific information in the context of the sample anatomy. Further, it fills the scale gap between the tomographic and microscopy techniques, imaging samples in the 1-20 mm size regime. A commercial OPT system, the Bioptonics 3001 OPT scanner was developed by Sharpe and over 30 units were sold worldwide [91] before the product and its support were discontinued in 2013 [90]. Most existing custom-built implementations are expensive, require specialized equipment, and do not include software for reconstruction. To enable a wider uptake of this technique and address the issues with cost and easy access to hardware and software, the following chapter describes the development and characterisation of a low-cost, open source OPT system for imaging large organ samples.

## References

- [1] G N Hounsfield. Computerized transverse axial scanning (tomography): Part 1. Description of system. *The British Journal of Radiology*, 46(552):1016–1022, dec 1973.
- [2] R. Damadian. Tumour Detection by Magnetic Resonance Imaging. *Science*, 171(3976):1151, 1971.
- [3] J C Elliott and S D Dover. X-ray microtomography. *Journal of Microscopy*, 126(2):211–213, 1982.
- [4] Dylan B. Smith, Galina Bernhardt, Nigel E. Raine, Richard L. Abel, Dan Sykes, Farah Ahmed, Inti Pedroso, and Richard J. Gill. Exploring miniature insect brains using micro-CT scanning techniques. *Scientific Reports*, 6(February):1–10, 2016.
- [5] Hrvoje Lusic and Mark W. Grinstaff. *X-ray-computed tomography contrast agents*, volume 113. 2013.
- [6] Che Wei Liao, Lih Jyh Fuh, Yen Wen Shen, Heng Li Huang, Chih Wei Kuo, Ming Tzu Tsai, and Jui Ting Hsu. Self-assembled micro-computed tomography for dental education. *PLoS ONE*, 13(12):1–11, 2018.

- [7] Armen R Kherlopian, Ting Song, Qi Duan, Mathew A Neimark, Ming J Po, John K Gohagan, and Andrew F Laine. A review of imaging techniques for systems biology. *BMC Systems Biology*, 2(1):74, 2008.
- [8] Vasilis Ntziachristos. Going deeper than microscopy : the optical imaging frontier in biology. *Nature Publishing Group*, 7(8):603–614, 2010.
- [9] Yoonsuck Choe. *Physical Sectioning Microscopy*, pages 1–4. Springer New York, New York, NY, 2013.
- [10] José Angel Conchello and Jeff W. Lichtman. Optical sectioning microscopy. *Nature Methods*, 2(12):920–931, 2005.
- [11] Wolfgang J. Weninger, Stefan H. Geyer, Timothy J. Mohun, Diego Rasskin-Gutman, Takaaki Matsui, Ines Ribeiro, Luciano Da F Costa, Juan Carlos Izpisúa-Belmonte, and Gerd B. Müller. High-resolution episcopic microscopy: A rapid technique for high detailed 3D analysis of gene activity in the context of tissue architecture and morphology. *Anatomy and Embryology*, 211(3):213–221, 2006.
- [12] Wolfgang J Weninger, Stefan Meng, Johannes Streicher, and Gerd B Müller. A new episcopic method for rapid 3-D reconstruction: Applications in anatomy and embryology. *Anatomy and Embryology*, 197(5):341–348, 1998.
- [13] Stefan H. Geyer, Barbara Maurer-Gesek, Lukas F. Reissig, and Wolfgang J. Weninger. High-resolution episcopic microscopy (HREM) - Simple and robust protocols for processing and visualizing organic materials. *Journal of Visualized Experiments*, 2017(125):1–11, 2017.
- [14] C. J.R. Sheppard and A. Choudhury. Image formation in the scanning microscope. *Optica Acta*, 24(10):1051–1073, 1977.
- [15] James B Pawley. *Handbook of Biological Confocal Microscopy*. New York, NY, third edition, 2006.
- [16] Jan Huiskens, Jim Swoger, Filippo Del Bene, Joachim Wittbrodt, and Ernst H K Stelzer. Optical Sectioning Deep Inside Live Embryos by Selective Plane Illumination Microscopy. *Science*, 305(5686):1007–1009, aug 2004.
- [17] P. J. Keller, A. D. Schmidt, J. Wittbrodt, and E. H.K. Stelzer. Reconstruction of Zebrafish Early Embryonic Development by Scanned Light Sheet Microscopy. *Science*, 322(5904):1065–1069, 2008.

- [18] Bi Chang Chen, Wesley R. Legant, Kai Wang, Lin Shao, Daniel E. Milkie, Michael W. Davidson, Chris Janetopoulos, Xufeng S. Wu, John A. Hammer, Zhe Liu, Brian P. English, Yuko Mimori-Kiyosue, Daniel P. Romero, Alex T. Ritter, Jennifer Lippincott-Schwartz, Lillian Fritz-Laylin, R. Dyche Mullins, Diana M. Mitchell, Joshua N. Benbenek, Anne Cecile Reymann, Ralph Böhme, Stephan W. Grill, Jennifer T. Wang, Geraldine Seydoux, U. Serdar Tulu, Daniel P. Kiehart, and Eric Betzig. Lattice light-sheet microscopy: Imaging molecules to embryos at high spatiotemporal resolution. *Science*, 346(6208), 2014.
- [19] Bo Jui Chang, Mark Kittisopikul, Kevin M. Dean, Philippe Roudot, Erik S. Welf, and Reto Fiolka. Universal light-sheet generation with field synthesis. *Nature Methods*, 16(3):235–238, 2019.
- [20] Raju Tomer, Khaled Khairy, Fernando Amat, and Philipp J Keller. Quantitative high-speed imaging of entire developing embryos with simultaneous multiview light-sheet microscopy. *Nature Methods*, 9(7):755–763, 2012.
- [21] Fabian F Voigt, Daniel Kirschenbaum, Evgenia Platonova, Stéphane Pagès, Robert A A Campbell, Rahel Kastli, Martina Schaettin, Ladan Egolf, Alexander Van Der Bourg, Philipp Bethge, Karen Haenraets, Noémie Frézel, Thomas Topilko, Paola Perin, Daniel Hillier, Sven Hildebrand, Anna Schueth, Alard Roebroek, Botond Roska, Esther T Stoeckli, Roberto Pizzala, Nicolas Renier, Hanns Ulrich Zeilhofer, Theofanis Karayannis, Urs Ziegler, Laura Batti, Anthony Holtmaat, Christian Lüscher, Adriano Aguzzi, and Fritjof Helmchen. The mesoSPIM initiative : open-source light-sheet microscopes for imaging cleared tissue. *Nature Methods*, 16(November), 2019.
- [22] D Huang, EA A Swanson, CP P Lin, J. S Schuman, W. G Stinson, W Chang, M. R Hee, T Flotte, K Gregory, C. A Puliafito, Al. Et, and Et al. Optical coherence tomography., 1991.
- [23] Silke Aumann, Sabine Donner, Jörg Fischer, and Frank Müller. Optical Coherence Tomography (OCT): Principle and Technical Realization. In *High Resolution Imaging in Microscopy and Ophthalmology*, pages 59–85. 2019.
- [24] Manmohan Singh, Raksha Raghunathan, Victor Piazza, Anjul M. Davis-Loiacono, Alex Cable, Tegy J. Vedakkan, Trevor Janecek, Michael V. Frazier, Achuth Nair, Chen Wu, Irina V. Larina, Mary E. Dickinson, and Kirill V. Larin. Applicability, usability, and limitations of murine embryonic imaging with optical coherence tomography and optical projection tomography. *Biomedical Optics Express*, 7(6):2295, 2016.



- [25] Judith Weber, Paul C Beard, and Sarah E Bohndiek. Contrast agents for molecular photoacoustic imaging. *Nature Publishing Group*, 13(8):639–650, 2016.
- [26] Paul Beard. Biomedical photoacoustic imaging. (June):602–631, 2011.
- [27] Amalina Binte, Ebrahim Attia, Ghayathri Balasundaram, Mohesh Moothanchery, U S Dinish, Renzhe Bi, Vasilis Ntziachristos, and Malini Olivo. A review of clinical photoacoustic imaging : Current and future trends. *Photoacoustics*, 16(November):100144, 2019.
- [28] Jelena Levi, Sri-rajashekar Kothapalli, Sarah Bohndiek, Joon-kee Yoon, Anca Dragulescu-andrasi, Carsten Nielsen, Aleksandra Tisma, Sunil Bodapati, Gayatri Gowrishankar, Xinrui Yan, Carmel Chan, Daniela Starcevic, and Sanjiv Sam Gambhir. Molecular Photoacoustic Imaging of Follicular Thyroid Carcinoma. *Clinical Cancer Research*, 19(6):1494–1503, 2013.
- [29] Sarah E Bohndiek, Laura S Sasportas, Steven Machtaler, Jesse V Jokerst, Sharon Hori, and Sanjiv S Gambhir. Photoacoustic Tomography Detects Early Vessel Regression Antiangiogenic Therapy Trebananib. *The Journal of Nuclear Medicine*, 56(12):1942–1948, 2015.
- [30] Werner Spalteholz. Über das Durchsichtigmachen von menschlichen und tierischen Präparaten. *Leipzig, Ger.: S. Hirzel*, 1914.
- [31] Hans-Ulrich Dodt, Ulrich Leischner, Anja Schierloh, Nina Jährling, Christoph Peter Mauch, Katrin Deininger, Jan Michael Deussing, Matthias Eder, Walter Zieglgänsberger, and Klaus Becker. Ultramicroscopy: three-dimensional visualization of neuronal networks in the whole mouse brain. *Nature Methods*, 4(4):331–336, 2007.
- [32] Ali Ertürk, Klaus Becker, Nina Jährling, Christoph P. Mauch, Caroline D. Hojer, Jackson G. Egen, Farida Hellal, Frank Bradke, Morgan Sheng, and Hans Ulrich Dodt. Three-dimensional imaging of solvent-cleared organs using 3DISCO. *Nature Protocols*, 7(11):1983–1995, 2012.
- [33] Kazuki Tainaka, Akihiro Kuno, Shimpei I. Kubota, Tatzuya Murakami, and Hiroki R. Ueda. *Chemical Principles in Tissue Clearing and Staining Protocols for Whole-Body Cell Profiling*, volume 32. 2016.
- [34] Ghayathri Balasundaram, Chris Jun, Hui Ho, Kai Li, Wouter Driessen, Chi Lok Wong, Vasilis Ntziachristos, Bin Liu, and Malini Olivo. Molecular photoacoustic imaging of

- breast cancer using an actively targeted conjugated polymer. *International Journal of Nanomedicine*, pages 387–397, 2015.
- [35] Kanyi Pu, Adam J Shuhendler, Jesse V Jokerst, Jianguo Mei, Sanjiv S Gambhir, Zhenan Bao, and Jianghong Rao. Semiconducting polymer nanoparticles as photoacoustic molecular imaging probes in living mice. *Nature Nanotechnology*, 9(3):233–239, 2014.
- [36] Adrian Taruttis and Vasilis Ntziachristos. Advances in real-time multispectral optoacoustic imaging and its applications. *Nature Photonics*, 9(4):219–227, 2015.
- [37] Lihong V. Wang and Song Hu. Photoacoustic tomography: In vivo imaging from organelles to organs. *Science*, 335(6075):1458–1462, 2012.
- [38] Gael Diot, Stephan Metz, Aurelia Noske, Evangelos Liapis, Barbara Schroeder, Saak V. Ovsepian, Reinhard Meier, Ernst Rummeny, and Vasilis Ntziachristos. Multispectral optoacoustic tomography (msot) of human breast cancer. *Clinical Cancer Research*, 23(22):6912–6922, 2017.
- [39] Rory M Power and Jan Huisken. A guide to light-sheet fluorescence microscopy for multiscale imaging. *Nature Methods*, 14(4):360–373, 2017.
- [40] Karen J.I. Lee, Grant M. Calder, Christopher R. Hindle, Jacob L. Newman, Simon N. Robinson, Jerome J.H.Y. Avondo, and Enrico S. Coen. Macro optical projection tomography for large scale 3D imaging of plant structures and gene activity. *Journal of Experimental Botany*, 68(3):527–538, 2017.
- [41] Pedro P. Vallejo Ramirez, Joseph Zammit, Oliver Vanderpoorten, Fergus Riche, Francois Xavier Blé, Xiao Hong Zhou, Bogdan Spiridon, Christopher Valentine, Simeon E. Spasov, Pelumi W. Oluwasanya, Gemma Goodfellow, Marcus J. Fantham, Omid Siddiqui, Farah Alimagham, Miranda Robbins, Andrew Stretton, Dimitrios Simatos, Oliver Haderler, Eric J. Rees, Florian Ströhl, Romain F. Laine, and Clemens F. Kaminski. OptiJ: Open-source optical projection tomography of large organ samples. *Scientific Reports*, 9(15693), 2019.
- [42] J. Sharpe. Optical Projection Tomography as a Tool for 3D Microscopy and Gene Expression Studies. *Science*, 296(5567):541–545, 2002.
- [43] A.C. Kak and Malcolm Slaney. *Principles of Computerized Tomography*. IEEE Press, 1988.

- [44] Ernst Abbe. Beiträge zur theorie des mikroskops und der mikroskopischen wahrnehmung. *Archiv für Mikroskopische Anatomie*, 9:413–418, 1873.
- [45] James Sharpe. Optical projection tomography as a new tool for studying embryo anatomy. *Journal of Anatomy*, 202(2):175–181, 2003.
- [46] Azriel Rosenfeld and A.C. Kak. *Digital Picture Processing*. Academic Press, second edition, 1982.
- [47] L A Shepp and B F Logan. The Fourier Reconstruction of a Head Section. *IEEE Transactions on Nuclear Science*, 21(L):21–43, 1974.
- [48] G. N. Ramachandran and A. V. Lakshminarayanan. Three-dimensional reconstruction from radiographs and electron micrographs: application of convolutions instead of Fourier transforms. *Proceedings of the National Academy of Sciences of the United States of America*, 68(9):2236–2240, 1971.
- [49] Martin J Willemink. The evolution of image reconstruction for CT — from filtered back projection to artificial intelligence. pages 2185–2195, 2019.
- [50] D. F. Swinehart. The Beer-Lambert Law. *Journal of Chemical Education*, 39(7):333, 1962.
- [51] Johnathon R. Walls, John G. Sled, James Sharpe, and R. Mark Henkelman. Correction of artefacts in optical projection tomography. *Physics in Medicine and Biology*, 50(19):4645–4665, 2005.
- [52] Lingling Chen, Sunil Kumar, Douglas Kelly, Natalie Andrews, Margaret J Dallman, Paul M W French, and James McGinty. Remote focal scanning optical projection tomography with an electrically tunable lens. *Biomedical optics express*, 5(10):3367–75, 2014.
- [53] Lingling Chen, James McGinty, Harriet B. Taylor, Laurence Bugeon, Jonathan R. Lamb, Margaret J. Dallman, and Paul M. W. French. Incorporation of an experimentally determined MTF for spatial frequency filtering and deconvolution during optical projection tomography reconstruction. *Optics Express*, 20(7):7323, 2012.
- [54] Teresa Correia, Nicola Lockwood, Sunil Kumar, Jun Yin, Marie Christine Ramel, Natalie Andrews, Matilda Katan, Laurence Bugeon, Margaret J. Dallman, James McGinty, Paul Frankel, Paul M W French, and Simon Arridge. Accelerated optical

- projection tomography applied to in vivo imaging of zebrafish. *PLoS ONE*, 10(8):1–17, 2015.
- [55] Samuel P.X. Davis, Sunil Kumar, Yuriy Alexandrov, Ajay Bhargava, Gabriela da Silva Xavier, Guy A. Rutter, Paul Frankel, Erik Sahai, Seth Flaxman, Paul M.W. French, and James McGinty. Convolutional neural networks for reconstruction of undersampled optical projection tomography data applied to in vivo imaging of zebrafish. *Journal of Biophotonics*, (August):1–10, 2019.
- [56] Emilio J Gualda, Tiago Vale, Pedro Almada, José A Feijó, Gabriel G Martins, and Nuno Moreno. OpenSpinMicroscopy: an open-source integrated microscopy platform. *Nature Methods*, 10(7):599–600, 2013.
- [57] Jelle van der Horst and Jeroen Kalkman. Image resolution and deconvolution in optical tomography. *Opt. Express*, 24(21):24460–24472, 2016.
- [58] Olli Koskela, Toni Montonen, Birhanu Belay, Edite Figueiras, Sampsa Pursiainen, and Jari Hyttinen. Gaussian Light Model in Brightfield Optical Projection Tomography. *Scientific Reports*, 9(1):1–14, 2019.
- [59] Jürgen Mayer, Alexandre Robert-Moreno, Renzo Danuser, Jens V. Stein, James Sharpe, and Jim Swoger. OPTiSPIM: integrating optical projection tomography in light sheet microscopy extends specimen characterization to nonfluorescent contrasts. *Optics letters*, 39(4), 2014.
- [60] Thomas Watson, Natalie Andrews, Samuel Davis, Laurence Bugeon, D Dallman, and James McGinty. OPTiM : Optical projection tomography integrated microscope using open-source hardware and software. *PLoS ONE*, pages 1–13, 2017.
- [61] M E Dickinson, A M Flenniken, X Ji, L Teboul, M D Wong, J K White, T F Meehan, W J Weninger, H Westerberg, H Adissu, C N Baker, L Bower, J M Brown, L B Caddle, F Chiani, D Clary, J Cleak, M J Daly, J M Denegre, B Doe, M E Dolan, S M Edie, H Fuchs, V Gailus-Durner, A Galli, A Gambadoro, J Gallegos, S Guo, N R Horner, C W Hsu, S J Johnson, S Kalaga, L C Keith, L Lanoue, T N Lawson, M Lek, M Mark, S Marschall, J Mason, M L McElwee, S Newbigging, L M Nutter, K A Peterson, R Ramirez-Solis, D J Rowland, E Ryder, K E Samocha, J R Seavitt, M Selloum, Z Szoke-Kovacs, M Tamura, A G Trainor, I Tudose, S Wakana, J Warren, O Wendling, D B West, L Wong, A Yoshiki, D G MacArthur, G P Tocchini-Valentini, X Gao, P Flicek, A Bradley, W C Skarnes, M J Justice, H E Parkinson, M Moore, S Wells,

- R E Braun, K L Svenson, M H de Angelis, Y Herault, T Mohun, A M Mallon, R M Henkelman, S D Brown, D J Adams, K C Lloyd, C McKerlie, A L Beaudet, M Bucan, and S A Murray. High-throughput discovery of novel developmental phenotypes. *Nature*, 537(7621):508–514, 2016.
- [62] José M Bioucas-Dias and Mário A T Figueiredo. A New TwIST: Two-Step Iterative Shrinkage/ Thresholding Algorithms for Image Restoration. *IEEE TRANSACTIONS ON IMAGE PROCESSING*, page 3, 2007.
- [63] Tomas Alanentalo, Andreas Hörnblad, Sofia Mayans, Anna Karin Nilsson, James Sharpe, Asa Larefalk, Ulf Ahlgren, and Dan Holmberg. Quantification and three-dimensional imaging of the insulinitis-induced destruction of beta-cells in murine type 1 diabetes. *Diabetes*, 59(7):1756–1764, 2010.
- [64] Gregory a Anderson, Michael D Wong, Jian Yang, and R Mark Henkelman. 3D imaging, registration, and analysis of the early mouse embryonic vasculature. *Developmental dynamics : an official publication of the American Association of Anatomists*, 242(5):527–38, 2013.
- [65] Jacqueline a. Gleave, Michael D. Wong, Jun Dazai, Maliha Altaf, R. Mark Henkelman, Jason P. Lerch, and Brian J. Nieman. Neuroanatomical phenotyping of the mouse brain with three-dimensional autofluorescence imaging. *Physiological Genomics*, 44(June 2012):778–785, 2012.
- [66] D Avid N Guyen, P A U L J M Archand, A L Rielle, P Lanchette, J Ulia N Ilsson, M Iguel S Ison, J Érôme, E Xtermann, A Ntonio L Opez, M Arcin S Ylwestrzak, J Jessica S Ordet Essimoz, A N J A S Chmidt Hristensen, D An, H Olmberg, D Imitri V A N D E V Ille, and T H E O L Asser. Optical projection tomography for rapid whole mouse brain imaging. *Biomedical Optics Express*, 8(12):331–336, 2017.
- [67] Andreas Hörnblad, Abbas Cheddad, and Ulf Ahlgren. An improved protocol for optical projection tomography imaging reveals lobular heterogeneities in pancreatic islet and  $\beta$ -cell mass distribution. *Islets*, 3(4):204–208, 2011.
- [68] Manuela Kellner, Marko Heidrich, Rebecca Beigel, Raoul-Amadeus Lorbeer, Lars Knudsen, Tammo Ripken, Alexander Heisterkamp, Heiko Meyer, Mark Philipp Kühnel, and Matthias Ochs. Imaging of the mouse lung with scanning laser optical tomography (SLOT). *Journal of applied physiology (Bethesda, Md. : 1985)*, 113(6):975–83, 2012.

- [69] Janet Kerwin, Mark Scott, James Sharpe, Luis Puellas, Stephen C Robson, Margaret Martínez-de-la Torre, Jose Luis Ferran, Guangjie Feng, Richard Baldock, Tom Strachan, Duncan Davidson, and Susan Lindsay. 3 Dimensional Modelling of Early Human Brain Development Using Optical Projection Tomography. *BMC neuroscience*, 5:27, 2004.
- [70] Johnathon R. Walls, Leigh Coultas, Janet Rossant, and R. Mark Henkelman. Three-dimensional analysis of vascular development in the mouse embryo. *PLoS ONE*, 3(8), 2008.
- [71] Natalie Andrews, Marie Christine Ramel, Sunil Kumar, Yuriy Alexandrov, Douglas J. Kelly, Sean C. Warren, Louise Kerry, Nicola Lockwood, Antonina Frolov, Paul Frankel, Laurence Bugeon, James McGinty, Margaret J. Dallman, and Paul M W French. Visualising apoptosis in live zebrafish using fluorescence lifetime imaging with optical projection tomography to map FRET biosensor activity in space and time. *Journal of Biophotonics*, 9(4):414–424, 2016.
- [72] A. Bassi, B. Schmid, and J. Huisken. Optical tomography complements light sheet microscopy for in toto imaging of zebrafish development. *Development*, 142(5):1016–1020, 2015.
- [73] Andrea Bassi, Luca Fieramonti, Cosimo D’Andrea, Marina Mione, Gianluca Valentini, and Cosimo D’Andrea. In vivo label-free three-dimensional imaging of zebrafish vasculature with optical projection tomography. *Journal of biomedical optics*, 16(10):100502, 2011.
- [74] Luca Fieramonti, Andrea Bassi, Efrem Alessandro Foglia, Anna Pistocchi, Cosimo D’Andrea, Gianluca Valentini, Rinaldo Cubeddu, Sandro de Silvestri, Giulio Cerullo, and Franco Cotelli. Time-Gated Optical Projection Tomography Allows Visualization of Adult Zebrafish Internal Structures. *PLoS ONE*, 7(11):1–7, 2012.
- [75] Sunil Kumar, Nicola Lockwood, Marie-Christine Ramel, Teresa Correia, Matthew Ellis, Yuriy Alexandrov, Natalie Andrews, Rachel Patel, Laurence Bugeon, Margaret J. Dallman, Sebastian Brandner, Simon Arridge, Matilda Katan, James McGinty, Paul Frankel, and Paul M. W. French. Quantitative in vivo optical tomography of cancer progression & vasculature development in adult zebrafish. *Oncotarget*, 5(28):2–11, 2016.
- [76] James McGinty, Khadija B. Tahir, Romain Laine, Clifford B. Talbot, Christopher Dunsby, Mark A A Neil, Laura Quintana, James Swoger, James Sharpe, and Paul M W

- French. Fluorescence lifetime optical projection tomography. *Journal of Biophotonics*, 1(5):390–394, 2008.
- [77] Alicia Arranz, Di Dong, Shouping Zhu, Charalambos Savakis, Jie Tian, and Jorge Ripoll. In-vivo optical tomography of small scattering specimens: Time-lapse 3D imaging of the head eversion process in *Drosophila melanogaster*. *Scientific Reports*, 4:1–5, 2014.
- [78] Leeanne McGurk, Harris Morrison, Liam P. Keegan, James Sharpe, and Mary A. O’Connell. Three-dimensional imaging of *Drosophila melanogaster*. *PLoS ONE*, 2(9), 2007.
- [79] Heiko Meyer, Alex Darrell, Athanasios Metaxakis, Charalambos Savakis, and Jorge Ripoll. Optical Projection Tomography for In-Vivo Imaging of *Drosophila melanogaster*. *Microscopy and Analysis*, 22(5):19–21, 2008.
- [80] Claudio Vinegoni, Chrysoula Pitsouli, Daniel Razansky, Norbert Perrimon, and Vasilis Ntziachristos. In vivo imaging of *Drosophila melanogaster* pupae with mesoscopic fluorescence tomography. *Nature Methods*, 5(1):45–47, 2008.
- [81] K. Lee, Jerome Avondo, Harris Morrison, Lilian Blot, Margaret Stark, James Sharpe, Andrew Bangham, and Enrico Coen. Visualizing Plant Development and Gene Expression in Three Dimensions Using Optical Projection Tomography. *the Plant Cell*, 18(9):2145–2156, 2006.
- [82] Matthias Rieckher, Udo Jochen Birk, Heiko Meyer, Jorge Ripoll, and Nektarios Tavernarakis. Microscopic optical projection tomography in vivo. *PLoS ONE*, 6(4):2–7, 2011.
- [83] Tomas Alanentalo, Amir Asayesh, Harris Morrison, Christina E. Lorén, Dan Holmberg, James Sharpe, and Ulf Ahlgren. Tomographic molecular imaging and 3D quantification within adult mouse organs. *Nature Methods*, 4(1):31–33, 2007.
- [84] Alicia Arranz, Di Dong, Shouping Zhu, Markus Rudin, Christos Tsatsanis, Jie Tian, and Jorge Ripoll. Helical optical projection tomography. *Optics Express*, 21(22):25912, 2013.
- [85] Varsha Kumar, Elke Scandella, Renzo Danuser, Lucas Onder, Maximilian Nitschké, Yoshinori Fukui, Cornelia Halin, Burkhard Ludewig, and Jens V. Stein. Global lymphoid tissue remodeling during a viral infection is orchestrated by a B cell-lymphotoxin-dependent pathway. *Blood*, 115(23):4725–4733, 2010.

- [86] Jean-François Colas and James Sharpe. Live optical projection tomography. *Organogenesis*, 5(4):211–6, 2009.
- [87] Edite Figueiras, Ana M Soto, Danilo Jesus, M Lehti, J Koivisto, J E Parraga, J Silva-Correia, J M Oliveira, R L Reis, M Kellomäki, and J Hyttinen. Optical projection tomography as a tool for 3D imaging of hydrogels. *Biomedical optics express*, 5(10):3443–3449, 2014.
- [88] Malcolm E. Fisher, Allyson K. Clelland, Andrew Bain, Richard A. Baldock, Paula Murphy, Helen Downie, Cheryll Tickle, Duncan R. Davidson, and Richard A. Buckland. Integrating technologies for comparing 3D gene expression domains in the developing chick limb. *Developmental Biology*, 317(1):13–23, 2008.
- [89] J McGinty, H B Taylor, L Chen, L Bugeon, J R Lamb, M J Dallman, and P M French. In vivo fluorescence lifetime optical projection tomography. *Biomed Opt Express*, 2(5):1340–1350, 2011.
- [90] Michael D. Wong, Jun Dazai, Johnathon R. Walls, Nicholas W. Gale, and R. Mark Henkelman. Design and Implementation of a Custom Built Optical Projection Tomography System. *PLoS ONE*, 8(9), 2013.
- [91] Invention, licensing and commercialisation of optical projection tomography microscopy. *Ref2014, Impact case study (REF3b)*, (Fall):14–15, 2009.

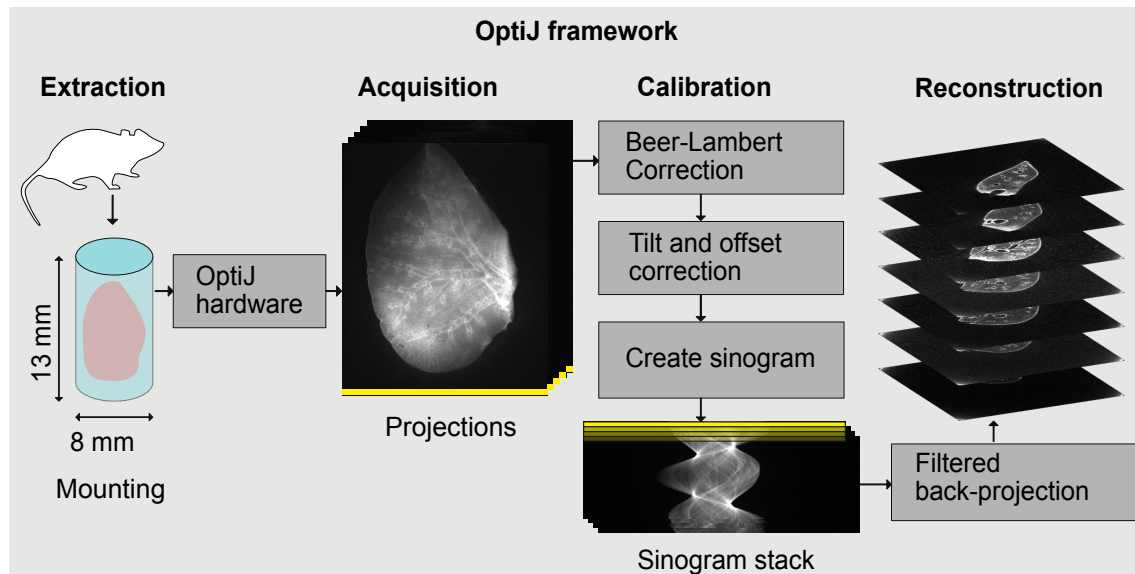


## Chapter 6

# Development of a low-cost, open-source OPT system to image large organ samples

This chapter describes the development, characterisation, and implementation of a low-cost, open-source OPT system, OptiJ, with which to image large organ samples.

OptiJ was designed to be accessible and easy to reproduce for other researchers. It was built using off-the-shelf optical components, and it includes an open-source ImageJ plugin library for OPT data reconstruction. The capabilities of the system were demonstrated by imaging full-sized adult mouse lungs that were cleared and immunostained; their study is relevant in the context of chronic obstructive pulmonary disease (COPD), which is a type of lung disease characterised by heterogeneously distributed emphysema (alveolar cell death) and bronchoconstriction (narrowing of airways). OptiJ allowed us to explore the morphology of the airway tree and visualise, in 3D, the tertiary airways, bronchioles, and alveolar sacs in mouse lung lobes. Fig. 6.1 summarises the end-to-end workflow of OptiJ, from sample preparation to imaging and reconstruction. This project was performed in collaboration with the 2015 cohort of the EPSRC Centre for Doctoral Training in Sensor Technologies and Applications (Sensor CDT), a group of doctoral students who conceived the idea of an open-source OPT system and built a working prototype, and two scientists from AstraZeneca who first devised the use of OPT to examine mouse lungs in a COPD model. My contribution to the project was to improve on both the hardware and software components of the prototype, characterise the system performance, and prepare, image, and reconstruct the adult mouse lung samples. Individual contributions to each portion of the work are specified accordingly. This work was published by Vallejo Ramirez et al. in [1], and this chapter is based on that publication.



**Fig. 6.1** Schematic showing the OptiJ workflow. An excised sample is embedded in agarose (Fig. 6.2) and mounted on a rotation stage; projections are then acquired over a full rotation of the sample using the hardware set up shown in Fig. 6.3. The OptiJ plugins are used for calibration and pre-processing, and subsequent 3D reconstruction of the sample (Fig. 6.4). Figure adapted from [1].

## 6.1 Motivation

The availability of open-source software and hardware tools allows the widespread adoption of technologies which are otherwise too expensive or unavailable for users worldwide. As discussed in Chapter 5, volumetric imaging techniques, including most existing OPT implementations, often require expensive equipment and technical expertise to build, and are therefore not easily accessible for many research labs. Furthermore, most techniques are limited to imaging samples up to a few millimetres in size. The motivation behind OptiJ was to enable a more widespread uptake of OPT, particularly for imaging large organs or organisms on the 5-20 mm scale. A fully-custom OPT has been previously described by Wong et al. [2], however its components are expensive (> £70,000) and it uses commercial software (LabView and Matlab) for its instrument control and reconstructions. To address the cost and technical complexity of the Wong et al. implementation, Watson et al. [3] developed an open-source adapter that fits any commercial inverted widefield microscope stage and allows for the sample alignment and rotation necessary for OPT. This modular implementation makes use of a microscope's existing components to perform OPT, therefore it is inexpensive and easy to implement. However, the sample size is limited by the plastic tubing used to hold the sample horizontally over the objective, which is at most two or

three millimetres in diameter [3]. OptiJ was inspired by the OpenSPIN microscope [4], a low-cost open-source OPT built with off-the-shelf components used to image small embryos. This implementation allowed for imaging large samples; however, it did not include its own software for image calibration and reconstruction (Bruker's nrecon software was used). To address the issues in previous open-source OPT implementations, a fully integrated OPT system with open-source software and hardware was developed specifically for imaging large tissue samples.

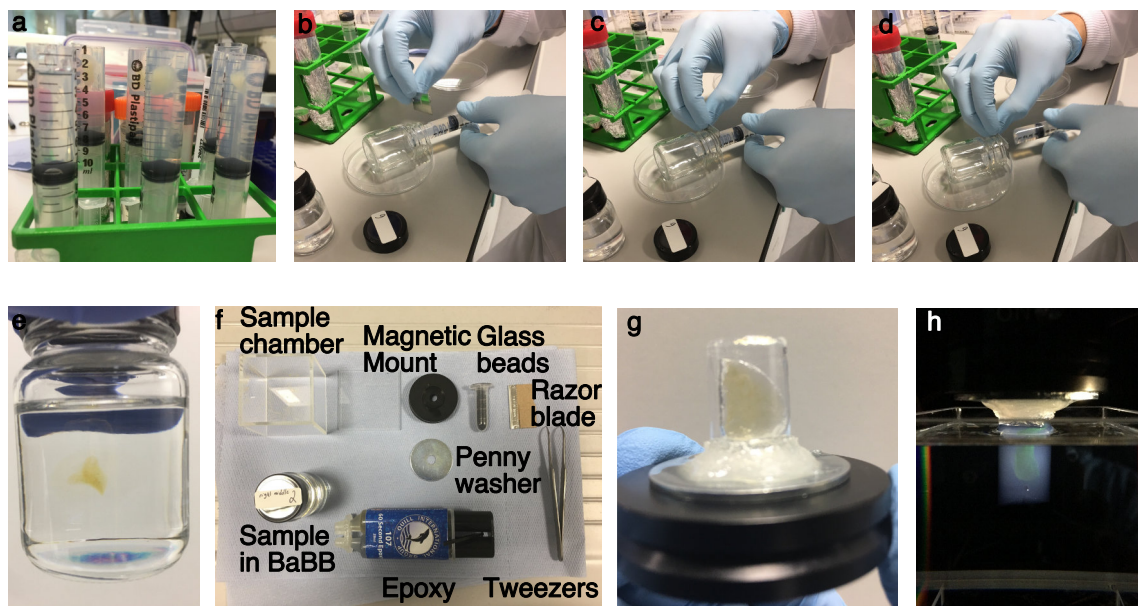
## **6.2 Materials and Methods**

### **6.2.1 Animal perfusion and tissue preparation**

For the immunostaining of the lungs, the mice were perfused through the right ventricle with phosphate-buffered saline (PBS) to remove blood from the tissue. The lungs were subsequently inflated with 4% PFA and fixed overnight at room temperature in fixative. Over the next three days, the lungs were rinsed in PBS and permeabilised through two cycles of dehydration-rehydration in a gradient of methanol, and in a solution of PBS and detergent (1% Triton X-100) to ensure antigens from the deepest part of the tissue were rendered accessible. All immunostains were then performed in 1% Triton X-100 in PBS (PBST) containing 10% of donkey serum. Two different immunostains were tested in separate lung samples with primary: i) anti-Surfactant C protein antibody to target membrane antigen secreted from airway type 2 epithelial cells in alveoli, or ii) anti-thyroid transcription factor-1 (TTF-1) antibody (Dako Agilent Products, mouse monoclonal, clone 8G7G3/1, Cat #M3575) to target nuclear antigen also present in airway type 2 epithelial cells. The lungs were incubated in primary antibody solution for one hour at room temperature and for 48 hours at 4°C, followed by extensive washes with PBST and 1% foetal calf serum. Fluorescent labelling of the primary antibody was achieved with anti-IgG Alexa Fluor 488 (AF488) secondary antibody in 1:500 dilution for 48 hours at 4°C followed by extensive washes with PBS for three hours to overnight. This procedure was performed by Dr. Xiao-Hong Zhou. A detailed version of the perfusion and immunostaining protocol can be found in OptiJ online repository: <https://lag-opt.github.io>. The stained and dehydrated lung samples were delivered to Cambridge, where my contribution to the project began for all subsequent sample clearing, mounting, imaging, and analysis steps.

### 6.2.2 Clearing and mounting whole lungs for imaging

Fixed and immunostained samples were embedded in a 2% low-melting-point agarose (Thermofisher Part #R0801) solution as a holding medium for sample clearing and image acquisition. 10 mL syringes were cut, using a razor blade, at the 1 mL and 6 mL mark. The syringe plunger was inserted from the 6 mL end just so the rubber tip was completely inside the cropped syringe tube; a pipette was then used to fill approximately three quarters of the available volume in the tube with molten agarose. The agarose was left to cool for 3-10 minutes before samples were carefully transferred into the agarose-filled tube using smooth tweezers, and were oriented close to the centre of the tube (Fig. 6.2 a).



**Fig. 6.2** Sample embedding and mounting for OPT imaging. (a-d) Embedding of the lung lobes in agarose. (e-h) Mounting of the cleared lobes on a magnetic stage ready for imaging.

A 1 mm diameter spherical glass bead (Sigma-Aldrich Part #Z250465-1PAK) was immediately inserted close to the sample, but not in the same horizontal plane to avoid occlusions, as a tracking fiducial for alignment and calibration during post-processing. The exposed end of the tube was sealed with parafilm to avoid dehydration of the agarose during storage. Samples were placed in a fridge at 4°C for one hour to allow the solution to fully cross-link into solid agarose cylinders. The embedded lung lobes were pushed out of the syringes (Fig. 6.2 b-d), and dehydrated using 50% methanol for 24 hours and 100% methanol for 48 hours. Embedded samples were subsequently immersed in a 1:2 mixture of Benzyl alcohol and Benzyl benzoate (BaBB) for 72 hours, changing the BaBB solution every 24 hours (Fig. 6.2 f). Prior to OPT acquisition, the agarose-embedded tissue cylinders were

glued onto bright-zinc plated (BZP) penny washers (M5x25, Fixmart Part #402203217) using quick-dry epoxy (Loctite Epoxy Quick Set 0.85-Fluid Ounce Syringe, Henkel Corporation, Part# 1395391). After the glue was cured, the penny washer was coupled to a magnetic kinematic mount (Thorlabs Part #SB1), ready to be inserted into the system for imaging (Fig. 6.2 g-h). This procedure was performed by me and Dr. Francois-Xavier Blé. A detailed step-by-step procedure for sample embedding and mounting can be found in the OptiJ online repository: <https://lag-opt.github.io>.

### 6.2.3 Experimental set-up

A 3D-printed flexure stage for open-source microscopy [5] was chosen for x,y,z translation and rotation of the sample because of its low cost (cost of printing material only, ~£5) and modular design. An Andor CLARA CCD camera with  $6.45 \times 6.45 \mu\text{m}^2$  pixels was initially used for acquisition of the OPT projections in this study, but a lower-cost camera can be implemented in a future release. A 0.5x telecentric objective (Edmund Optics Part #63-741) with a 65 mm working distance and 0.028 NA was chosen to acquire the maximum field of view possible with the camera. Two white light LEDs (Thorlabs Part #MWWHD3) were chosen to provide even illumination with minimal flicker in both tOPT and eOPT. These were fitted in small cage systems with an optical diffuser (Thorlabs Part #DG10-600), an adjustable iris (Thorlabs #SM1D12D), and a condenser lens (Thorlabs Part #LA1401-A). A GFP excitation and emission filter pair was used for eOPT (Excitation: 482/25 Part #FF01-482/25-25, Emission: 515/LP Part #FF01-515/LP-25, Semrock). A Hellma glass cuvette (Z805750-1EA, Scientific Laboratory Supplies) was used as the immersion chamber for the sample during imaging (Fig. 6.2 h). The filled chamber was raised using a Swiss Boy lab jack (Sigma-Aldrich Part #2635316-1EA) to completely cover the agarose gel containing the sample during the acquisitions with the immersion media.

The acquisition software was written in Java by Pelumi Oluwasanya from the Sensor CDT and packaged as an independent executable file. eOPT and tOPT projections were acquired with exposure times of 300 ms and 1 ms, respectively. The integration time and powers used for the fluorescence acquisition typically facilitated the capturing of signals of which the maximum and a mean signal correspond to ~20% and ~1.5% of the camera's full well capacity, respectively. Pictures of the set-up (Fig. 6.3), a list of parts, instructions for assembly, information about the acquisition software, and the characterization of the x, y, and z motion of the stage can be found in the OptiJ online repository: <https://lag-opt.github.io>. I received a functional version of the OPT hardware prototype and acquisition software and used it for all experiments shown in this chapter.

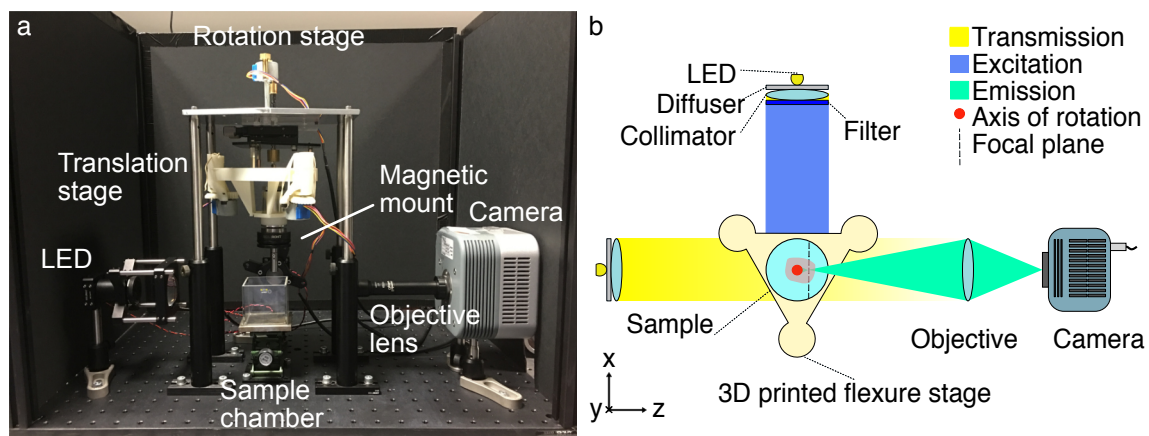
## 6.3 Characterisation of the software and hardware components

### 6.3.1 OptiJ hardware

OPT relies on the rotation of a sample to acquire 2D projections at different angles. As discussed in Chapter 5, assuming the thickness of the sample is smaller than the depth of field of the system, projections acquired over half a revolution are theoretically sufficient to recover an accurate 3D reconstruction of the sample structure. However, acquisitions over a full rotation typically lead to higher quality reconstructions [6, 7]. To minimise hardware errors prone to cause reconstruction artefacts, the OptiJ hardware design focused on the following considerations:

1. Ensuring the axis of rotation is parallel to the imaging plane of the camera,
2. Aligning the sample to the field of view of the camera
3. Robustly and repeatably performing the rotation of the sample and acquisition of the projections.

The OptiJ hardware enables the mounting, alignment, and rotation of thick biological samples for the acquisition of 2D projections in both eOPT and tOPT modalities. Figure 6.1 a shows the implemented set-up, which includes a monolithic 3D-printed, rotation and translation stage; a telecentric objective lens; a camera; two broadband LEDs; fluorescence excitation and emission filters; and collimating and diffusing optics.



**Fig. 6.3** OptiJ hardware set up. (a) Picture of the set up with labels on the main components. (b) Top-view illustration of the OptiJ hardware. Figure adapted from [1].

The main criteria guiding the component choices were widespread availability and low cost. The 3D-printed stage was adapted from the open-source Flexscope design [5] to accomplish the movement necessary for both linear alignment and rotation of the sample with stepper motors (Part #28BYJ-48). These motors were chosen specifically for their compatibility with the open-source stage and their low cost. The downside of using a low-cost motor for sample rotation is the presence of a larger amount of jitter (Fig. 6.4 b iii) than that attainable with more expensive motors. Two approaches were adopted to mitigate this drawback: the use of a bearing system to compensate for mechanical jitter from the motor, and a post-acquisition jitter correction during the image processing stage.

The stage achieves sub-micron steps in the x, y, and z direction (Fig. 6.3 b), with a maximal hysteresis of 58  $\mu\text{m}$  over a 3 mm travel range, and a maximum of 1024 steps per motor revolution corresponding to a minimum angular step size of  $0.35^\circ$ . The hysteresis was characterised by tethering a metal pin to the bottom of the translation stage, and subsequently moving the pin laterally across the FOV. This lateral motion was detected on the camera and the leading edge of the pin was tracked through multiple scans across the FOV. The tracked position of the pin was used to calculate the hysteresis over the travel range of the stage; the details on the stage characterisation can be found in [1] or in the OptiJ repository <https://lag-opt.github.io>. This characterisation was performed by Oliver Vanderpoorten and Fergus Riche. In practice, the center of the stage was aligned to the center of the FOV only once, and the stage was not translated again for any further experiments. A telecentric lens with low NA and 0.5x magnification was chosen to match the typical volume of adult mouse lungs. The low NA allows a depth of field of  $\sim 4$  mm, which upon sample rotation, allows for a maximum field of view (FOV) of  $13 \times 8 \times 8 \text{ mm}^3$ . The focal plane of the objective was placed midway between the axis of rotation and the front face of the sample such that only one half of the sample was in focus at any projection angle (as shown with the dashed line in Fig. 6.3 c). The telecentricity of the lens allowed us to use FBP for reconstruction, as only approximately parallel rays from the sample reached the camera. White light LEDs were chosen for their brightness, long life, and low-cost, and a custom circuit board was designed to minimise output flicker. The LED output was homogenised and collimated with off-the-shelf optics to ensure uniform illumination across the FOV.

The stage, the camera, and the LEDs were controlled with a Raspberry Pi 2B, which connected via USB to a computer which controlled the acquisition process and stored the captured images. Alternative choices for motors or stages to accomplish the rotation required in OPT, and for objective lenses compatible with OptiJ are provided in the OptiJ repository <https://lag-opt.github.io> or in [1]. This allows future users to choose which components best suit their application based on their desired cost to performance ratio, and the desired

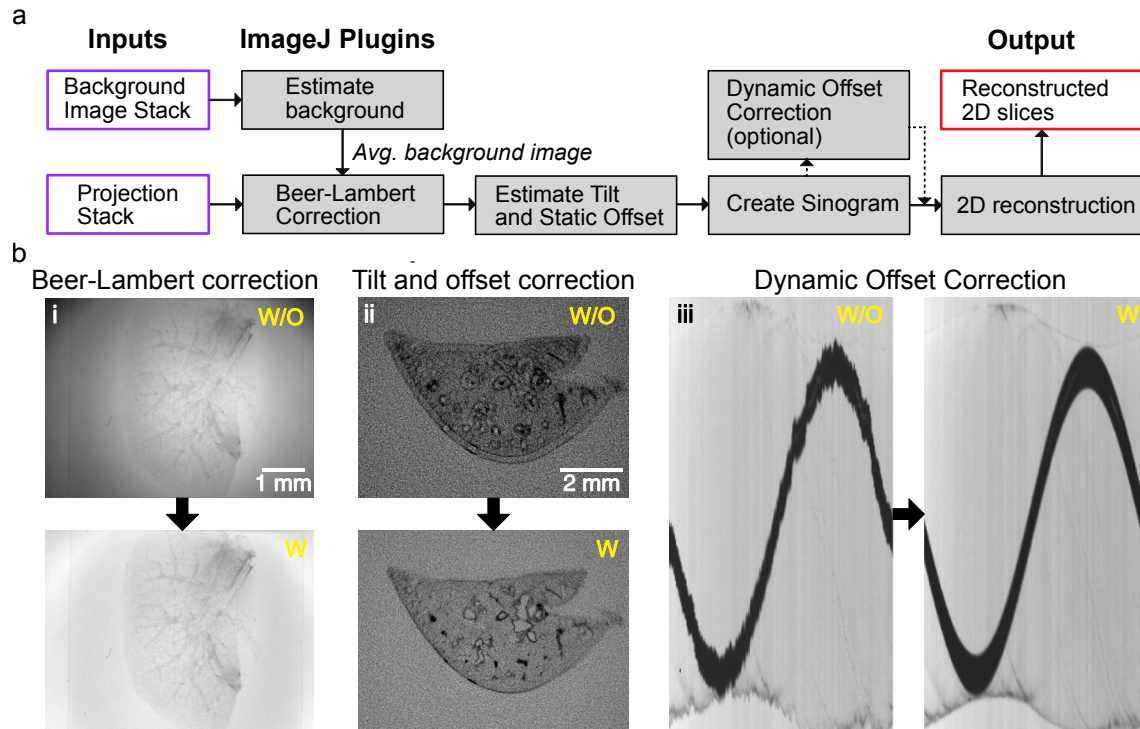
specifications of their system, e.g. magnification (and resulting FOV), resolution, working distance, and sample thickness. The hardware set up was first built by the Sensor CDT student cohort as a proof-of-concept project in 12 weeks. I characterised, rebuilt, and improved the OptiJ hardware by redesigning the collimation assembly for the two illumination arms, adding a telecentric objective, and adding fluorescence filters. This allowed for fluorescence imaging (eOPT), and more uniform illumination over a larger FOV.

### 6.3.2 OptiJ software

The reconstruction of high-quality 3D volumes from OPT projections requires data pre-processing to avoid artefacts during FBP, as discussed in Chapter 5 Section 5.2.1. To mitigate the effect of artefacts, a library of open-source ImageJ/Fiji plugins to pre-process OPT data were developed, as well as a GPU-accelerated FBP algorithm for reconstruction [8]. A suggested workflow for using the plugins is shown in Fig. 6.4, along with examples of corrected OPT artefacts.

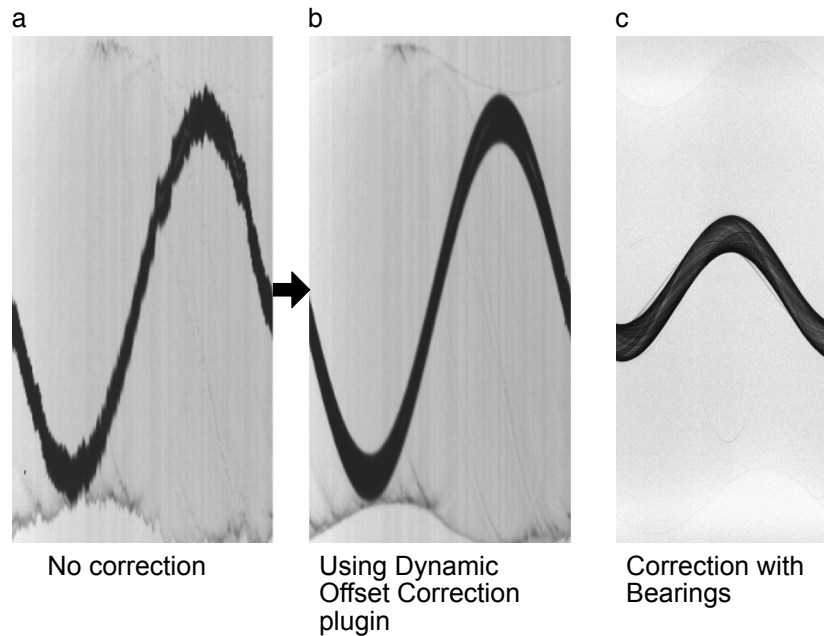
The *Beer-Lambert correction* plugin calibrates each tOPT projection using an average bright field image to obtain linear attenuation coefficients corrected for non-uniform pixel intensities, following the Beer-Lambert Law [9] and demonstrated in Fig. 6.4 b.i. A background image stack without the sample is acquired to obtain an average bright field image, and then each projection is divided by the average background image. Typical attenuation factors between dark and bright regions as shown in Fig. 6.4 b.i. range between two and three. A common artefact in OPT arises from the axis of rotation of the sample not being parallel to the plane of the FOV during acquisitions, which leads to the appearance of a shadow artefact around sharp features (Fig. 6.4 b.ii). The *Estimate Tilt and Offset* plugin tracks a fiducial marker (such as a 100  $\mu\text{m}$  glass bead) in the projections to determine if the axis of rotation is parallel to the plane of the FOV and produces correction values for the projection stack if this condition is not satisfied. These values can be used at the reconstruction step to minimise any shadow artefacts, as demonstrated in the corrected image in Fig. 6.4 b.ii. The *Create Sinogram* plugin displays a Radon Transform (sinogram) of the projections and uses the correction values for tilt and offset produced by the previous plugin to account for residual deviations, relaxing the need for a precise alignment of the system prior to acquisitions. Small sample wobble caused by mechanical jitter from low-cost stepper motors without a bearing system can be observed as jagged edges in an otherwise smooth sinogram (Fig. 6.4 b.iii). The *Dynamic Offset Correction* plugin calculates a sinusoidal fit of the motion of a fiducial marker and uses the difference between the ideal fit coordinates and the actual motion of the bead to produce a jitter-free sinogram as shown in the corrected image in Fig. 6.4 b.iii. This correction shown in the sinograms (Fig. 6.4 a-b) is similar to that obtained





**Fig. 6.4** OptiJ plugin library workflow for the correction of common OPT artefacts. (a) Suggested workflow for the use of the OptiJ plugins. (b) Correction of common OPT artefacts using OptiJ plugins. The top row represents images without correction applied (w/o). The bottom row shows images after correction (w). (i) Uneven illumination in raw tOPT projections resulting from the optics used to collimate the light source, and absorption and scattering from the sample. (ii) shadow artefact originating from a misalignment of the sample's axis of rotation. (iii) jittered sinogram of a marker bead rotated by a low-cost stepper motor. Figure adapted from [1].

using a high-quality bearing system that produces a smooth rotation of the sample, as shown in Fig. 6.5.



**Fig. 6.5** Comparison of Dynamic Offset Correction plugin and bearing system to compensate for mechanical jitter from low-cost motors. (a) Sinogram from raw projections of a glass bead in agarose showing jitter, visualised as jagged edges. (b) Correction of the jittered sinogram using the Dynamic Offset Correction plugin. (c) Correction of jitter using a high-quality bearing system coupled to the rotation motor and the sample holder.

This step concludes the pre-processing required to minimise artefacts in the reconstruction. The *2D reconstruction* plugin implements FBP to reconstruct a 3D cross-sectional stack of the original object using the corrected sinogram. To improve reconstruction times via FBP, the plugin allows for GPU-enabled acceleration using OpenCL [10], which is an open-source, platform-independent framework for parallel computing. This plugin also allows the user to choose from a variety of filters (Ramp, Hamming, Shepp-Logan, or no filter) for back-projection [11]. The OptiJ plugin library improves upon a previous implementation of the Radon transform in ImageJ/Fiji, which did not include calibration nor accelerated reconstruction algorithms [12]. A detailed description of the OptiJ plugin library, its functions and methods, usage and sample data for testing can be found in the OptiJ online repository at <https://lag-opt.github.io>. The code for the plugins was written and packaged by Joseph Zammit from the Sensor CDT cohort. I received a working version of the software and applied in this thesis for the reconstruction of the mouse lung data. My responsibility in this project began with the clearing and mounting of samples and included all subsequent imaging experiments and analyses performed.

## 6.4 Results

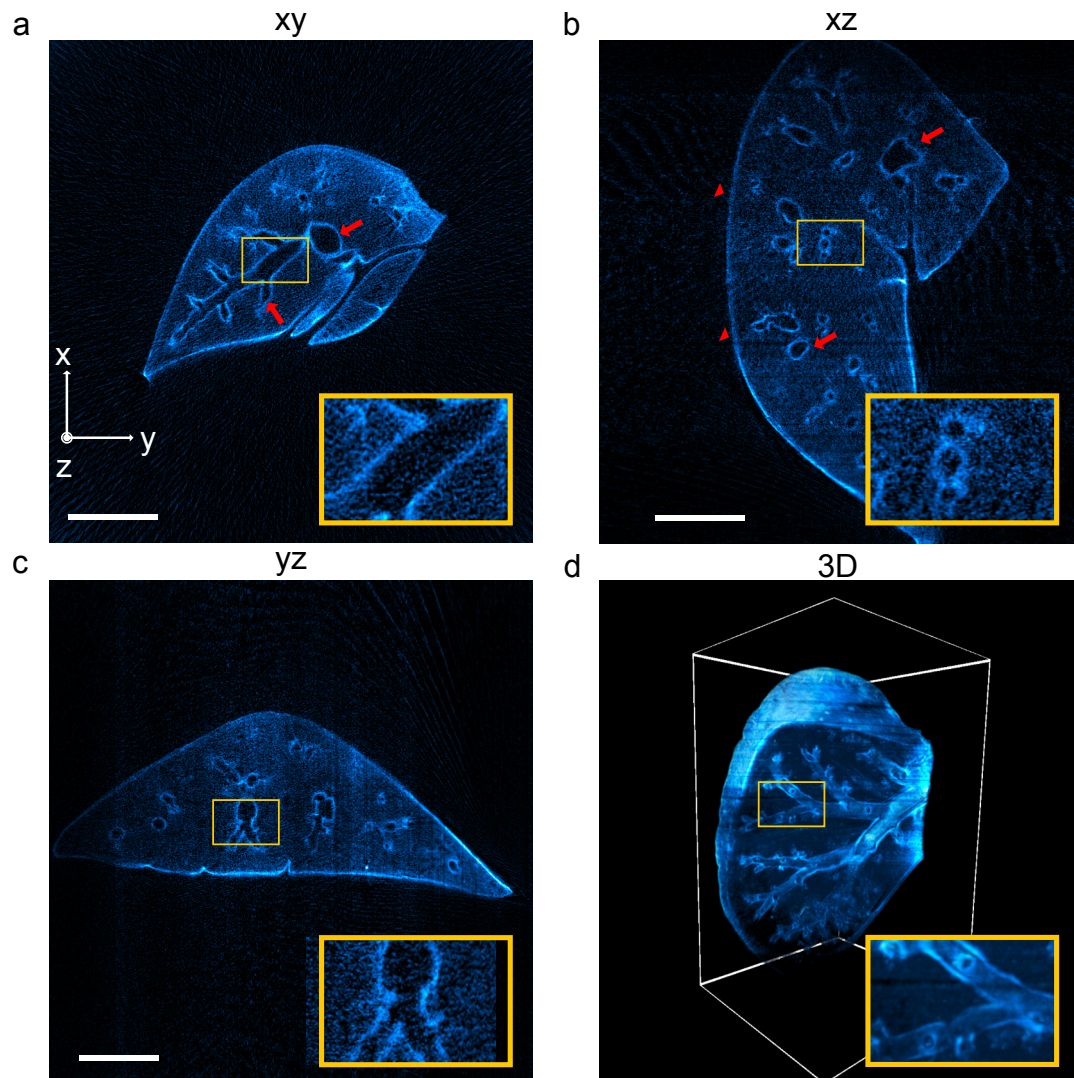
The non-destructive 3D imaging of whole lung lobes can be very useful in the study of COPD models in mice, as it allows for the identification of characteristic phenotypes such as bronchoconstriction (narrowing of airways), and the investigation of the extent of the structures affected in different lung areas. The OptiJ hardware and software tools described were applied to image fluorescently labelled and cleared mouse lungs. Two different antibodies were used to tag epithelial cells, and the resulting contrast and visible feature size with each label were analysed. The resolution achieved after OPT reconstruction was characterised using Fourier Ring Correlation (FRC), and the results obtained were corroborated using standard resolution measurements.

### 6.4.1 Visualisation of primary and higher order airways in different mouse lung lobes using eOPT

The superior, medial, and accessory lobes of the right lung, and the entire left lung of two adult mice were imaged using the OptiJ framework. 512 raw projections were acquired over a full rotation of each lobe to obtain high-fidelity reconstructions, as fewer projections lead to streak artefacts which degrade image quality and worsen the post-reconstruction resolution of the sample (Fig. 6.11). To obtain a morphological readout of the structures critical to studying COPD, such as the bronchioles and alveolar sacs, two different immunolabelling strategies which targeted proteins expressed in lung epithelial type 2 cells were tested and compared. In both strategies, a secondary antibody conjugated with an AF488 dye was used to visualise the airway tree through eOPT.

#### Surfactant protein C labelling

The lobes of the first mouse lung were immunostained with a primary antibody targeting the Surfactant protein C (Fig. 6.6 a-d). This labelling strategy revealed only gross features in the lobes' eOPT reconstructions, as demonstrated in the orthogonal views of the reconstructed stack from a large left lobe in Fig. 6.6 a-d. The primary bronchus and some secondary and tertiary airways are indicated by red arrows in Fig. 6.6 a-b, and the region in which the indiscernible finer features would be located, the parenchyma (lobe edge), is indicated by red arrowheads. The fluorescent signal collected with this labelling strategy likely originated from a combination of tissue autofluorescence, mostly from collagen, non-specific labelling, and the specific fluorescent signal from the structure of interest.



**Fig. 6.6** eOPT reconstructions of a left lobe labelled with anti-Surfactant Protein C AF488. Orthogonal views, xy (a), xz (b), and yz (c) and a 3D rendering (d) are shown. (a-d) The red arrows and insets indicate the primary airways visualised in the orthogonal cross-sections. The 3D rendering in (d) displays a clipping plane through the lung, highlighting secondary and tertiary bronchi in the inset. Interactive 3D renderings of the reconstructions can be found at <https://lag-opt.github.io/>. The acquisition time for this data set was 20 minutes, which includes a two second delay between each projection to avoid motion artefacts from the wobbling agarose gel. Figure adapted from [1]. Scale bars = 2 mm

### TTF-1 protein labelling

The lobes from the second mouse were immunostained with a primary antibody against the thyroid transcription factor-1 (TTF-1, see Fig. 6.7 a-d). In comparison to the label targeting Surfactant protein C, the labelling strategy targeting the TTF-1 protein produced significantly higher quality reconstructions which allowed for the visualisation of both large airways (~200-300  $\mu\text{m}$  in diameter) and minute bronchioles (~60-100  $\mu\text{m}$  in diameter) through the centre and periphery of the lobes. The orthogonal views of the reconstructed stack from a medial lobe show both the primary and secondary bronchi (red arrows in Fig. 6.7 a-b) and the higher order airways and tiny air sacs in the parenchyma (red arrowheads in Fig. 6.7 a-b). Figure 6.7 d shows a 3D rendering of the entire medial lobe; a zoomed inset directs attention to the intricate network of higher order airways that can be visualised inside the volume.

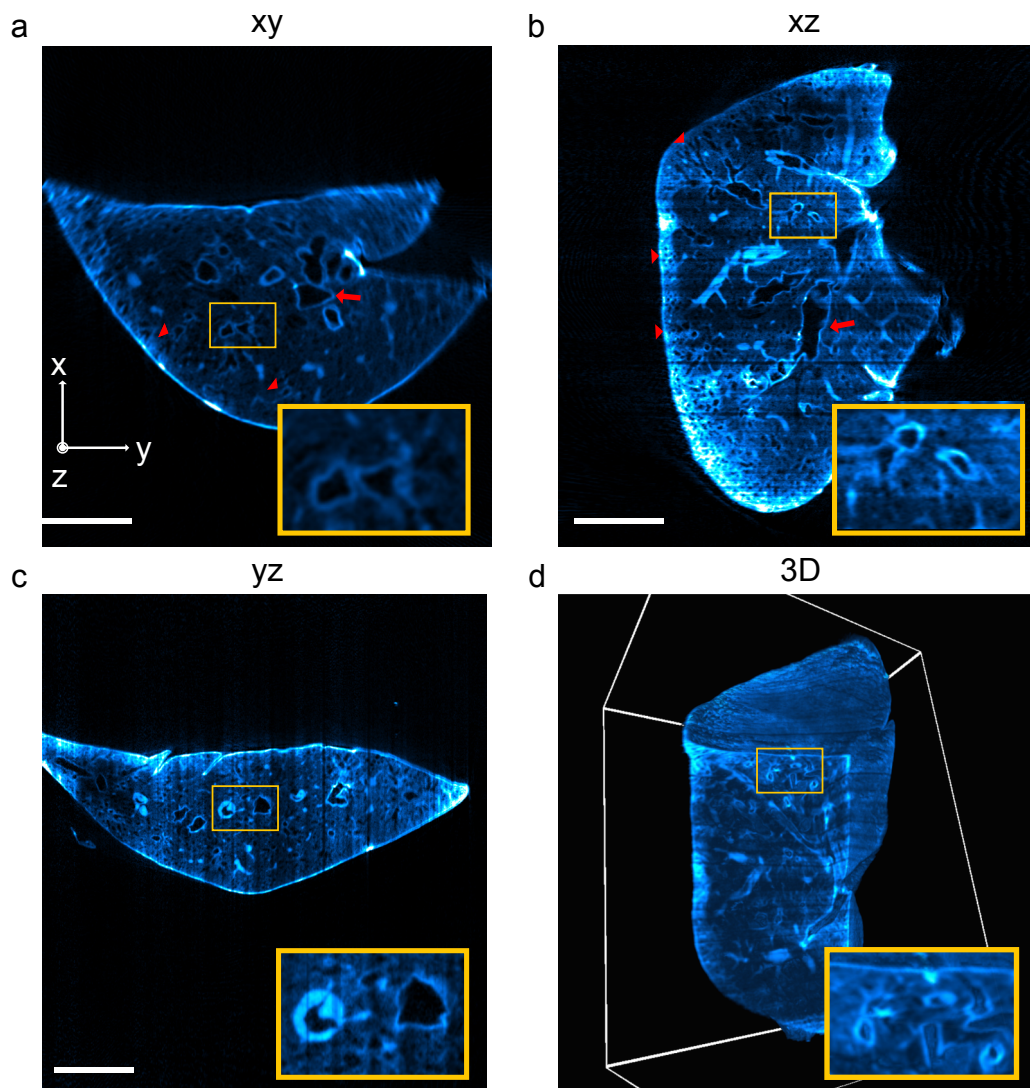
#### 6.4.2 Brightness and signal-to-noise ratio comparison between the two labelling strategies

To compare the post-reconstruction image quality resulting from the two labelling strategies, and to characterise the effect of non-specific labelling and autofluorescence on the total signal, we estimated the brightness ratio (BR) and the signal-to-noise ratio (SNR) between specifically labelled and background regions in the lung lobes shown in Fig. 6.6 and 6.7. Representative slices were chosen from the reconstructed stacks for the anti-surfactant C-labelled lobe (Fig. 6.8 a) and the anti-TTF1-labelled lobe (Fig. 6.8 b) and the ratio of specific fluorescent signal and non-specific signal was measured in two ways. The mean fluorescence intensity in Fig. 6.8 a-b was calculated in regions in which the signal was expected to be highly specific, such as alveolar cell clusters in higher order airways (yellow inset), and regions in which epithelial cells were less abundant and therefore the signal was expected to originate predominantly from autofluorescence and non-specific labelling (red inset). The mean intensities for regions exhibiting specific ( $I_1$ ) and non-specific ( $I_2$ ) signal were used to calculate the signal-to-noise ratio (SNR), using the standard deviation of the intensity measurements ( $\sigma$ ) as an estimate for the noise as defined below:

$$SNR = \frac{I_1 - I_2}{\sigma} \quad (6.1)$$

The resulting SNR for the anti-TTF-1 lobe was ~4, and that for the anti-Surfactant C lobe was ~1.6. Similarly, the brightness ratio (BR) was calculated between regions with signal and with background, as defined below:





**Fig. 6.7** eOPT reconstructions of a medial right lobe labelled with anti-TTF-1 AF488. Orthogonal views, xy (a), xz (b), and yz (c) and a 3D rendering (d) are shown. (a-d) The red arrows indicate a set of main airways (secondary and tertiary bronchi) in the medial lobe, and red arrowheads indicate high-order airways inside or close to the parenchyma. Small airways close to the primary bronchi are highlighted in the insets in (a) and (b). The 3D rendering in (d) with a clipping plane on one of the lobe faces shows a thick meshwork of higher order airways (quaternary bronchi and bronchioles). Interactive 3D renderings of the reconstructions can be found at <https://lag-opt.github.io>. The acquisition time for this data set was 20 minutes, which includes a two second delay between each projection to avoid motion artefacts from the wobbling agarose gel. Figure adapted from [1]. Scale bars = 2 mm

$$BR = \frac{I_1}{I_2} \quad (6.2)$$

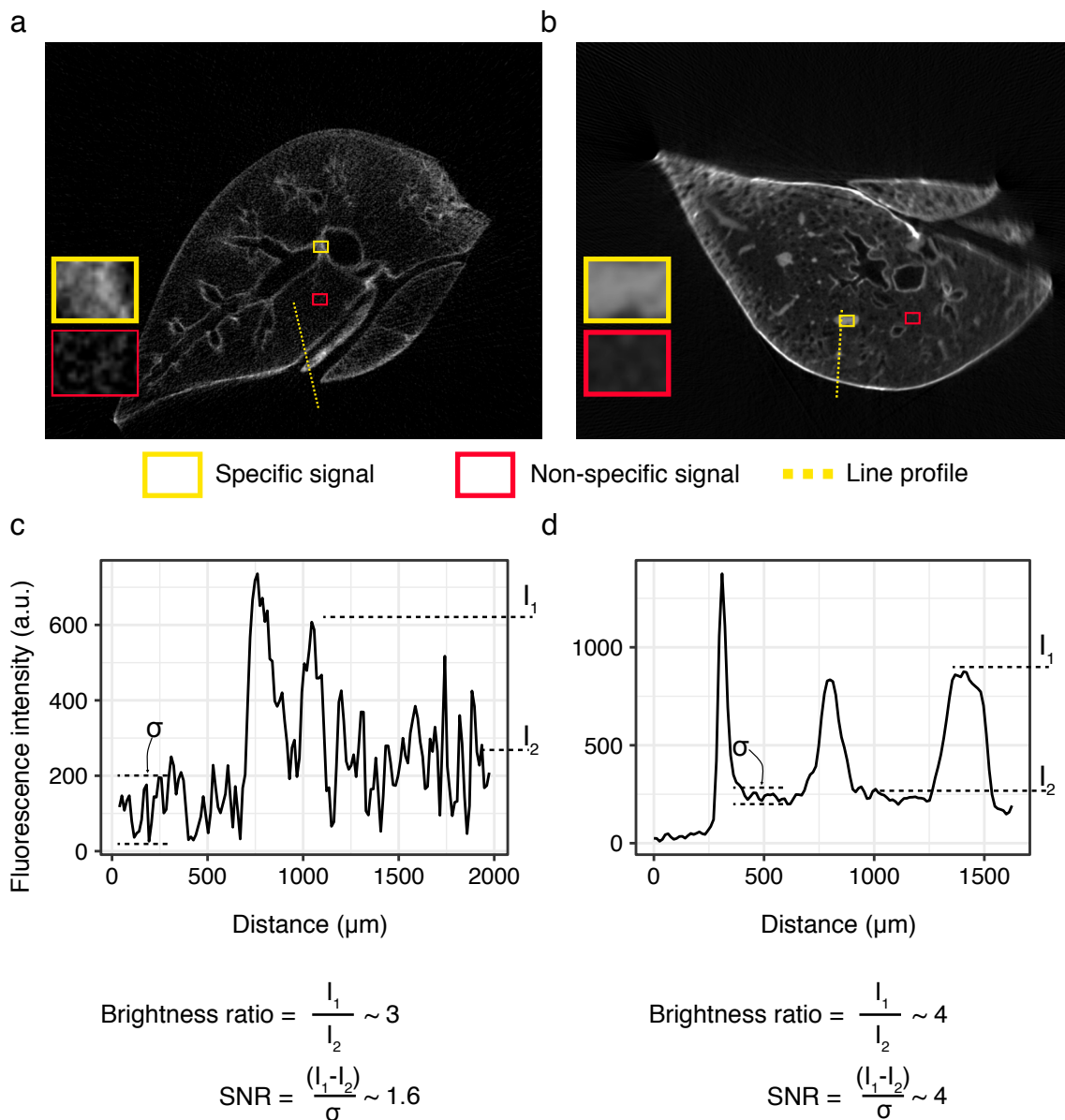
The resulting BR for the anti-TTF-1 lobe was  $\sim 4$  and that for the anti-surfactant protein C lobe was  $\sim 3$ . From this analysis, the TTF-1 labelling was shown to provide a 2.5-fold increase in SNR compared to Surfactant protein C. Additionally, the BR suggests that the background contributed between 25% and 33% of the total signal for both the TTF-1 and the Surfactant C labelling, respectively. However, no further repeats were analysed, therefore a more comprehensive study would be necessary to determine the effectiveness of these two labelling strategies for imaging epithelial tissue in whole lungs.

### 6.4.3 Histogram matching for volumetric data visualisation

During reconstruction via FBP of eOPT data, sharp edges or small features in the sample may cause streak artefacts in the reconstruction, particularly if these features exhibit a strong fluorescent signal. These streak artefacts result in very bright or dim cross-sectional slices in the reconstructions, as indicated by red arrows in the volumetric visualisation of the anti-Surfactant C left lung in Fig. 6.9 a. These are similar but not equal to the streak artefacts resulting from under sampling Radon space during reconstructions mentioned in Chapter 5, as these appear even with densely-sampled data sets such as the mouse lungs explored here. A histogram matching algorithm was implemented to account for these intensity variations, and visualisation without excessively bright or dim cross-sectional slices were obtained (Fig. 6.9 b). Essentially, the histogram matching algorithm reads in the volume data slice per slice, identifies the modal group in each slice's histogram, and shifts the modal peak to the origin. This reduces the discrepancies in brightness among slices and allows us to choose appropriate contrast thresholds to visualise the full volume.

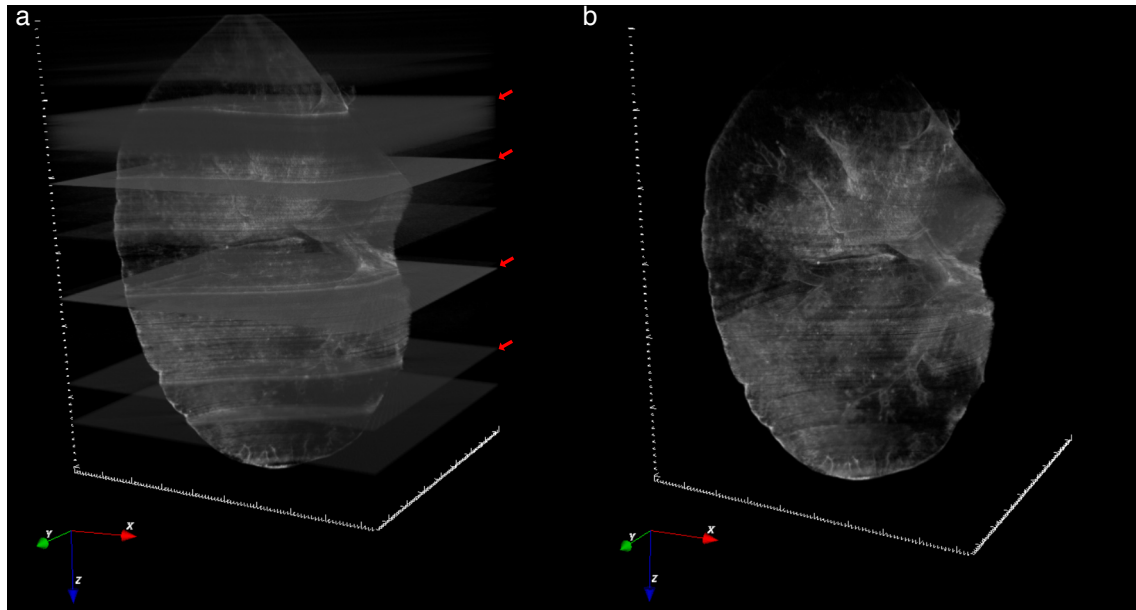
### 6.4.4 Quantifying resolution using Fourier Ring Correlation

The resolution in OPT reconstructions depends on the optical resolution dictated by the imaging lens with which the raw data were collected. It is further affected by the number of projections, as an insufficient number of projections leads to image artefacts (such as streaks) and thus degrades image quality. FRC is a well-established standard to measure resolution in electron microscopy [13], and more recently SMLM [14], as introduced briefly in Chapter 2 in Part I of this thesis. FRC analyses the similarity of two independent reconstructions of the same object in spatial frequency space to determine the feature size threshold at which both reconstructions are consistent. This threshold provides an estimate for the resolution limit of



**Fig. 6.8** Brightness and signal-to-noise ratio comparison between the two epithelial cell labelling strategies for mouse lungs. Representative slices from both the anti-Surfactant protein C labelled left lobe (a) and the anti-TTF-1 labelled right medial lobe (b) with yellow insets representing regions of specific signal and red insets representing regions with non-specific signal. (c-d) Representative line intensity profiles from (a) and (b) with labels on the mean intensities from regions in which the signal is specific ( $I_1$ ) and non-specific ( $I_2$ ), where the standard deviation is  $\sigma$ . Figure adapted from [1].



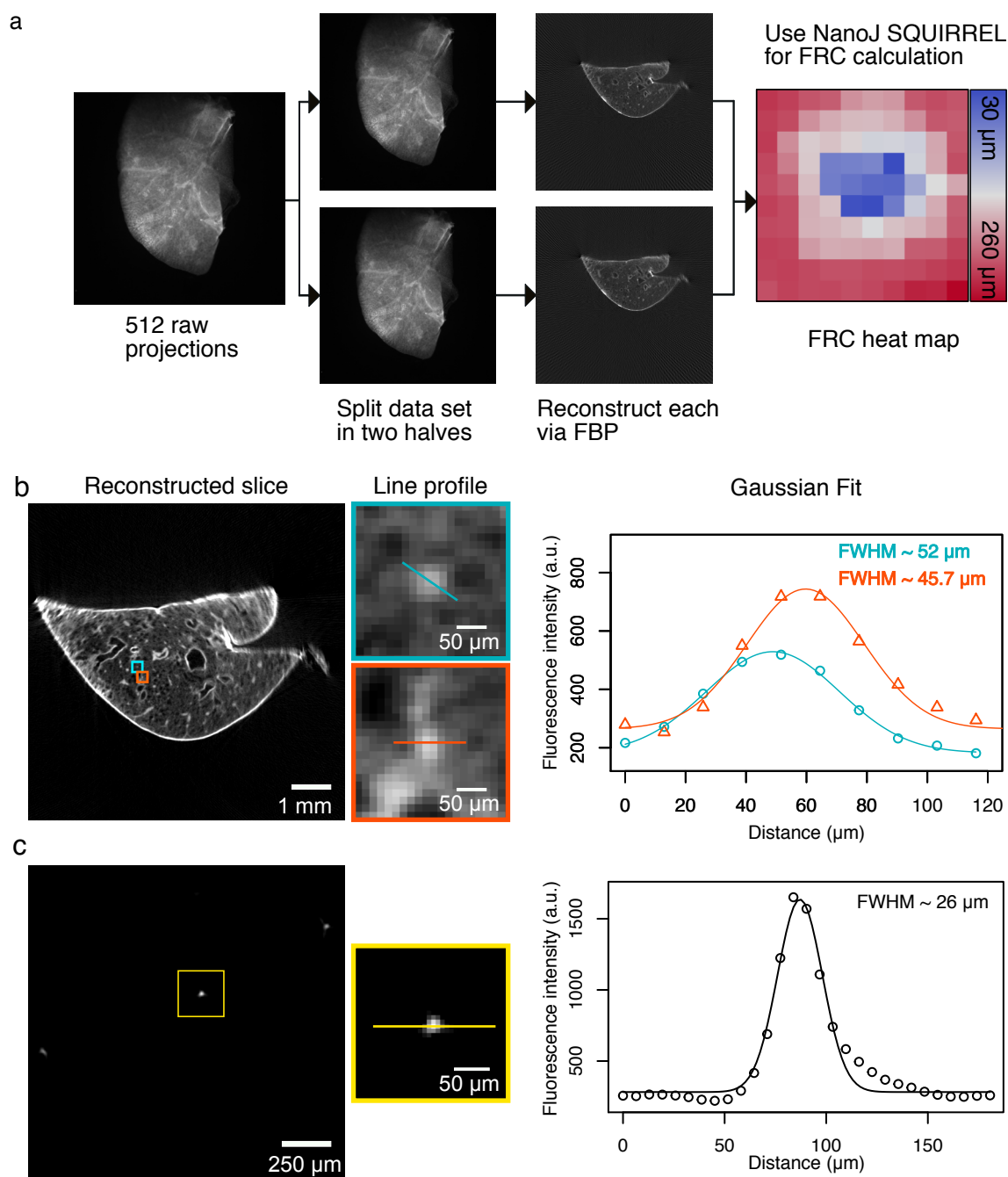


**Fig. 6.9** Histogram matching of cross-sectional slices is useful for volumetric visualisation of lung data. (a) eOPT reconstruction of anti-Surfactant C left lung in which bright or dim cross-sectional slices after FBP are highlighted with red arrowheads. (b) The whole lung volume after histogram matching.

the image. FRC was implemented here to measure post-reconstruction resolution in OPT (Fig. 6.10 a) with the medial lobe labelled with anti-TTF-1 AF488.

The raw data set with 512 projections was split into even and odd slice stacks; each half was reconstructed separately using the OptiJ *2D reconstruction* plugin to simulate two independent reconstructions of the same object. 20 cross-sectional slices were chosen in the middle of the sample, and the FRC calculation was performed using the NanoJ-SQUIRREL FRC plugin [15], resulting in a mean FRC value of  $49 \pm 16 \mu\text{m}$ . This FRC value and its uncertainty were obtained by averaging over the centre portion (blue squares) of the FRC heatmaps output by the plugin (Fig. 6.10 a), representing the pixel area occupied by the reconstructed lung cross-section. The default number of 10 spatial blocks per axis was used in SQUIRREL.

This novel approach to estimate resolution in OPT was compared to two standard resolution estimation methods: measuring cross-sections of fine sample features (Fig. 6.10 b) and measuring the cross-section of fluorescent beads (Fig. 6.10 c). The cross-sectional profiles in both cases were fitted to Gaussian functions, and the full width at half-maximum (FWHM) for the Gaussians was calculated as an estimate of the finest detail discernible. The FWHMs for the fine lung features in an arbitrary slice of the right medial lobe reconstruction in Fig. 6.10 b were  $\sim 52 \mu\text{m}$  (bottom) and  $\sim 46 \mu\text{m}$  (top). The FWHM for the fluorescent bead



**Fig. 6.10** Characterisation of OptiJ post-reconstruction resolution using three different methods. (a) Resolution estimate using FRC results in a resolution of  $49 \pm 16 \mu\text{m}$ . (b) Resolution estimate obtained by fitting Gaussian functions to the intensity cross sections of fine features in an arbitrary slice of the right medial lobe reconstruction. The FWHM for each Gaussian fit provides a measure for the size of the finest discernible features: (top) FWHM  $\sim 52 \mu\text{m}$ , (bottom) FWHM  $\sim 46 \mu\text{m}$  (c) Resolution estimate using fluorescent beads: a Gaussian function was fitted to the intensity cross section of a  $15 \mu\text{m}$ -diameter fluorescent bead reconstructed using OptiJ, yielding a FWHM of  $\sim 26 \mu\text{m}$ .

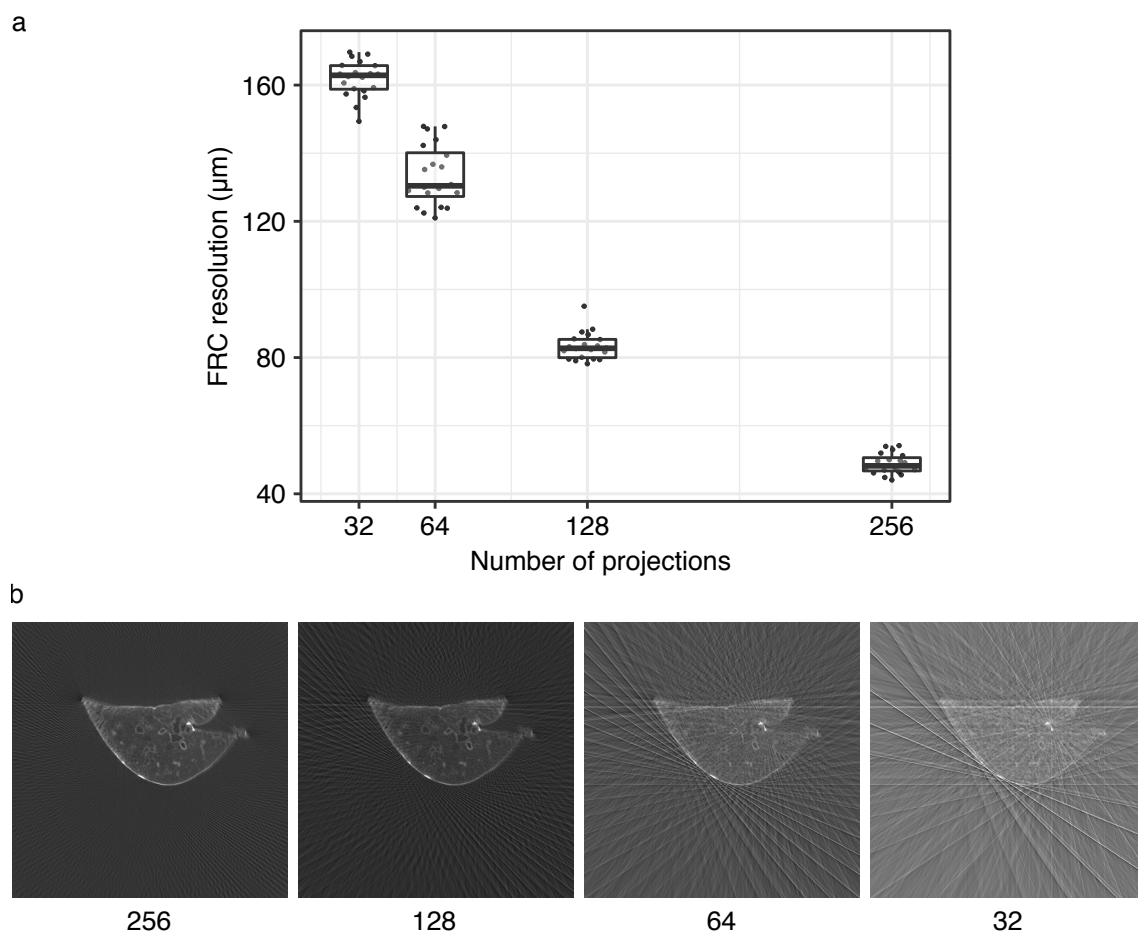
cross-section was  $\sim 26\text{ }\mu\text{m}$ . 256 projections of  $15\text{ }\mu\text{m}$ -diameter fluorescent beads immersed in agarose were acquired and reconstructed using OptiJ for the latter measurement.

The resolution estimates obtained with both methods are consistent with the output from FRC, confirming FRC is a reliable metric for estimating resolution in OPT. The FRC value for resolution is computed over the entire reconstructed slice, therefore it does not depend on user-drawn line profiles, and it provides an unbiased, quantitative estimate of resolution for OPT data sets. However, by effectively down sampling the data set from 512 to 256 projections to perform the independent object reconstructions for FRC, the resulting FRC value is an underestimate of the resolution of the original data set. This is due both to the decrease in the total signal contribution for any slice and therefore a higher noise component post-reconstruction, and to the streak artefacts resulting from having fewer projections for reconstruction. The effect of the number of projections used for OPT reconstruction on the FRC resolution estimate was quantified and illustrated in Fig. 6.11. The data set for the right medial lobe with 512 projections was down sampled into data sets with 256, 128, 64 and 32 projections and the FRC resolution estimate described in Fig. 6.10 was applied to 20 representative slices. Fewer projections lead to severe streak artefacts (Fig. 6.11 b) which degrade image quality and worsen the post-reconstruction resolution of the data set. The artefacts from the data splitting procedure are a major disadvantage of this method to evaluate resolution in OPT, compared to the more traditional methods of using sub-diffractive fluorescent beads which can provide an unbiased estimate of the true post-reconstruction lateral resolution. However, FRC does not require imaging an agarose phantom with embedded fluorescent beads and can therefore be considered a "calibration-free" method to evaluate resolution in already existing OPT data sets. Another advantage of FRC over traditional bead measuring methods is that FRC evaluates resolution over each lung slice for the entire volume and can therefore provide a resolution estimate as a function of distance away from the rotation axis, unlike bead phantoms in which fluorescent beads are sparsely distributed and therefore only sample a sparse subsection of the imaging available volume.

Considering the estimates from all three methods, in the presence of noise and sample jitter the post-reconstruction resolution of the system is likely to be between 25 and  $50\text{ }\mu\text{m}$ .

## 6.5 Summary and Discussion

This chapter described a low-cost, open-source hardware and software implementation of OPT for the examination of large volumetric samples. The imaging of whole organs in 3D with OptiJ at  $\sim 50\text{ }\mu\text{m}$  resolution was demonstrated. The method revealed the structure of



**Fig. 6.11** FRC resolution as a function of the number of projections acquired. (a) Plot of the FRC resolution estimate as a function of the number of projections available. (b) Representative images showing a reconstructed slice for each of the down-sampled data sets.

adult murine lungs, from the large primary bronchi to the minute bronchioles at the lung periphery. An open-source toolbox for OPT reconstructions, including a novel workflow for artefact correction and volume reconstruction was developed all in ImageJ, along with detailed characterisation and instructions for use.

The hardware challenges introduced by low-cost OPT solutions were addressed. In particular, the sensitivity to sample alignment was compensated for by tracking a marker glass bead and correcting for the estimated sample tilt using the OptiJ plugins provided. The jitter introduced by low-cost stepper motors used for sample rotation was addressed using two strategies. In one strategy, a bearing system was introduced to improve rotational stability on the motor used for sample rotation. In the other strategy, a novel software-based correction method was used, the *Dynamic Offset Correction* plugin, which can correct for jitter from low-cost motors in cases where high-quality motors or bearings are unaffordable or difficult to source. These measures ensure both accuracy and repeatability in the recording of high-fidelity OPT data. Furthermore, Fourier Ring Correlation (FRC) was implemented for the first time as a resolution measure for reconstructed OPT data sets. Consistent results were obtained with FRC compared to standard practices for measuring resolution.

The 3D imaging of intact lungs from COPD mouse models could be useful to visualise alveolar cell clusters in a whole-organ perspective, where the involvement of specific cell types in pathophysiological processes could be tracked and quantified. This could be a complementary approach to recent studies of COPD pathophysiology with confocal microscopy [16]. The immunostaining against the Surfactant protein C led to a diffuse signal, indicating a low specificity of the labelling in this context, potentially due to a deterioration of the Surfactant protein C by fixation, dehydration, or detergent treatment of the lobes during preparation. Nonetheless, the autofluorescence from elastin and collagen in epithelial cells, and extracellular matrix from the large airway wall likely boosted the fluorescent signal, such that images of the large airway tree with good contrast were obtained (Fig. 6.6). The anti-TTF-1 labelling proved more effective to visualise small airways with a better SNR compared to the Surfactant protein C label, especially in the lobe edges (Fig. 6.7). More generally, the 3D imaging data of intact mouse organs enabled with OptiJ could be useful in tracking specific cell types, visualising the heterogeneous distribution of disease, or assessing the effects of therapeutics in animal models of respiratory diseases.

In summary, a complete set of calibration and reconstruction routines in ImageJ/Fiji was described, along with an easy to build hardware set up for a complete OPT system. OptiJ implements both CPU and GPU acceleration for reconstructions, which yields reconstructions in tens of minutes rather than multiple hours. This is a key feature, as reconstruction of the substantial data sets resulting from imaging large organ samples is highly time-consuming

when performed sequentially with no acceleration. Furthermore, a larger field of view ( $13 \times 8 \times 8 \text{ mm}^3$ ) than most other OPT implementations [17, 18, 19, 20, 3] was demonstrated, which typically range from  $1 \times 1 \times 1 \text{ mm}^3$  to  $5 \times 5 \times 5 \text{ mm}^3$ . The larger field of view of OptiJ could be useful for future studies examining 3D anatomical structures and fluorescent signals from large model organisms, organ samples from small animals, or even organoids grown from pluripotent stem cells.

## 6.6 Future work

OptiJ is still in early stages of its development, and improvements to its hardware and software components would make it easier to adopt and use. Future work on the hardware could include automation of the tilt and offset calibration routines with a direct feedback loop to the hardware after correction with the OptiJ plugins. Alternatively, a simpler manual stage for the sample with a more accurate and stable motor for rotation could be used instead. A list of alternative motor choices for OPT rotation can be found in the Supplementary Information in [1]. Sample alignment is generally performed at the start of an imaging session, and minimal re-tuning is required during experiments so a manual stage would be adequate.

Improvements on software are perhaps more pressing, as the OptiJ plugins are readily available for researchers with existing tomography data sets and can be adopted faster than the hardware components. To speed up reconstructions, the OptiJ GPU acceleration kernel could be optimised and improved using new OpenCL libraries. To improve reconstruction quality, the deconvolution model proposed by Van der Horst [21] could be implemented to reduce the effect of the spatially-variant PSF. Another option would be to use the Gaussian-beam model for FBP reconstructions from Koskela et al [22] to provide higher quality reconstructions with better contrast. To reduce acquisition times, a compressed sensing approach as proposed by Correia et al. [6] could be adopted to allow for sparser data sets, and could be reconstructed with sufficient quality using only Fiji plugins. Alternatively, the use of convolutional neural networks to remove streak artefacts from sparsely sampled data could be applied to increase the throughput of the system as demonstrated by Davis et al. [23]. In order to improve the analysis of lung volumes reconstructed with OPT, an algorithm could be developed to quantify the physical volume of the detected airways as a way to evaluate the degree of bronchoconstriction, one of the hallmarks of COPD. This could be accomplished by thresholding areas with low contrast within the lung (empty airways) and adding their enclosed areas across all lung slices. Another algorithm could be developed to characterise the number and distribution of alveolar sacs at the lung parenchyma as a way to diagnose lung health, since healthy lungs will have fewer collapsed or destroyed air sacs

compared to diseased lungs. A future study using a higher NA objective lens during imaging, and therefore better resolution, could be performed to quantify the relative distribution of alveolar cell clusters in small lung regions for control vs COPD lungs, as a way to assess disease spread in the lung, and to perhaps evaluate the effect of drugs on alveolar cell recovery. New algorithms for the analysis of 3D lung imaging data, whether obtained with OPT or with other 3D imaging modalities [24, 25] could be of special interest to the scientific community for studies of COVID19, which primarily affects the lungs, and leaves significant lung damage post recovery. Regarding the lung imaging experiments, sample preparation could be improved by implementing newer tissue-clearing methods such as 3DISCO [26] and CLARITY [27] instead of the current approach based on BaBB, which is known to introduce loss of fluorescent signal from certain dyes [28] and may cause linear shrinking of tissue [29]. In terms of labelling, far-red dyes could be implemented in the future instead of AF488 for deeper penetration into tissue. In future experiments, it would be useful to image cleared, unlabelled lung lobes to properly characterise the signal contribution from autofluorescence in different lung regions. This could not only help our choice of fluorescent label for visualising small structures, but also it would allow us to better understand the available dynamic range of our system.

## 6.7 Perspectives

In order to harness the advantages of OPT and integrate this technique easily into existing imaging workflows in other laboratories, it is key to develop new tools to easily acquire, reconstruct, and analyse OPT data. The OptiJ hardware included useful tools such as the bearing system for sample wobble correction, and the Flexscope stage for sample alignment and automation of the acquisition procedure. However, building a specialised set up for OPT requires time and some engineering expertise to troubleshoot. A more widespread access of the technique could be achieved using a much simpler sample rotation system that can be readily adapted to existing widefield microscopes, similar to Watson et al.'s OPTiM [3]. Scientists working with large transparent biological samples (zebrafish, cleared organs, or cleared plants) could acquire OPT data with minimal modifications to their existing widefield microscopes, and subsequently process and reconstruct their data using OptiJ's open-source ImageJ plugins. Alternatively, OPT could be combined with expansion microscopy [30], a recently reported technique in which a fluorescently-labelled sample is physically enlarged, thereby increasing the effective resolution in a imaging system. In expansion microscopy, a specimen is embedded in a fine polymer matrix which covalently anchors any fluorescent labels within the specimen. Labels located closer than the

diffraction limit can then be isotropically separated and optically resolved with a conventional microscope. The physical structure of the specimen is destroyed; however, the relative location of the labelled fluorescent species remains intact, allowing previously unresolvable structures to be discerned. Expansion factors of 4 [30], 10 [31], 16 [32] have been reported; larger expansion factors remain available by iteratively expanding the sample at the cost of reduced fluorescent signal per sample volume [32]. By implementing expansion microscopy on a sample prior to imaging with OptiJ, a post-reconstruction resolution of up to 7 (4X expansion) and 2 (16X expansion)  $\mu\text{m}$  could be achieved. Challenges in gel mounting and sample labelling could need to be addressed, however this would represent a novel imaging modality which could be useful for imaging whole tissue slices in 3D with high resolution.

## 6.8 Acknowledgement of contributions

This project was a collaborative effort between the Sensor CDT 2015 cohort, two scientists from AstraZeneca, and the Laser Analytics Group. The Sensor CDT cohort delivered a functional prototype of the OPT system in October 2016 when I joined the project. Specifically, the OptiJ hardware was prototyped by Fergus Riche, Gemma Goodfellow, Omid Siddiqui, Tristan Hughes, and Dimitrios Simatos. The OptiJ software was written by Joseph Zammit, Bogdan Spiridon, and Simeon Spasov. The image acquisition software was written and compiled by Bogdan Spiridon and Pelumi Oluwasanya. Early experiments on bead samples were performed by Miranda Robbins, Oliver Vanderpoorten, Farah Alimagham, and Andrew Stretton. The Sensor CDT cohort efforts were coordinated by Christopher Valentine. The perfused and stained lung samples were provided by Dr. Xiao-Hong Zhou from the IMED Biotech Unit at AstraZeneca, Gothenburg, and the clearing and mounting of the samples was performed in collaboration with Dr. Francois-Xavier Blé from the IMED Biotech Unit at AstraZeneca, Cambridge. The detailed contributions of all specialists involved in this project can be found in the Acknowledgments section in [1]. Dr. Romain Laine and Dr. Florian Ströhl were instrumental in providing guidance and advice during all stages of the project. In particular, Dr. Romain Laine was steadfast in his advice on how to conduct experiments and analysis on the mouse lung samples, as well as in his guidance on writing the manuscript for publication. This work also benefitted from useful advice and guidance from Prof. Clemens Kaminski, Dr. Oliver Hader, Dr. Thomas Watson, and Dr. James McGinty.



## References

- [1] Pedro P. Vallejo Ramirez, Joseph Zammit, Oliver Vanderpoorten, Fergus Riche, Francois Xavier Blé, Xiao Hong Zhou, Bogdan Spiridon, Christopher Valentine, Simeon E. Spasov, Pelumi W. Oluwasanya, Gemma Goodfellow, Marcus J. Fantham, Omid Siddiqui, Farah Alimagham, Miranda Robbins, Andrew Stretton, Dimitrios Simatos, Oliver Haderler, Eric J. Rees, Florian Ströhl, Romain F. Laine, and Clemens F. Kaminski. OptiJ: Open-source optical projection tomography of large organ samples. *Scientific Reports*, 9(15693), 2019.
- [2] Michael D. Wong, Jun Dazai, Johnathon R. Walls, Nicholas W. Gale, and R. Mark Henkelman. Design and Implementation of a Custom Built Optical Projection Tomography System. *PLoS ONE*, 8(9), 2013.
- [3] Thomas Watson, Natalie Andrews, Samuel Davis, Laurence Bugeon, D Dallman, and James McGinty. OPTiM : Optical projection tomography integrated microscope using open-source hardware and software. *PLoS ONE*, pages 1–13, 2017.
- [4] Emilio J Gualda, Tiago Vale, Pedro Almada, José A Feijó, Gabriel G Martins, and Nuno Moreno. OpenSpinMicroscopy: an open-source integrated microscopy platform. *Nature Methods*, 10(7):599–600, 2013.
- [5] James P. Sharkey, Darryl C W Foo, Alexandre Kabla, Jeremy J. Baumberg, and Richard W. Bowman. A one-piece 3D printed flexure translation stage for open-source microscopy. *Review of Scientific Instruments*, 87(2), 2016.
- [6] Teresa Correia, Nicola Lockwood, Sunil Kumar, Jun Yin, Marie Christine Ramel, Natalie Andrews, Matilda Katan, Laurence Bugeon, Margaret J. Dallman, James McGinty, Paul Frankel, Paul M W French, and Simon Arridge. Accelerated optical projection tomography applied to in vivo imaging of zebrafish. *PLoS ONE*, 10(8):1–17, 2015.
- [7] J. Sharpe. Optical Projection Tomography as a Tool for 3D Microscopy and Gene Expression Studies. *Science*, 296(5567):541–545, 2002.
- [8] Johannes Schindelin, Ignacio Arganda-Carreras, Erwin Frise, Verena Kaynig, Mark Longair, Tobias Pietzsch, Stephan Preibisch, Curtis Rueden, Stephan Saalfeld, Benjamin Schmid, Jean-Yves Tinevez, Daniel James White, Volker Hartenstein, Kevin Eliceiri, Pavel Tomancak, and Albert Cardona. Fiji: an open-source platform for biological-image analysis. *Nat Meth*, 9(7):676–682, jul 2012.

- 
- [9] D. F. Swinehart. The Beer-Lambert Law. *Journal of Chemical Education*, 39(7):333, 1962.
- [10] John E Stone, David Gohara, and Guochun Shi. OpenCL: A Parallel Programming Standard for Heterogeneous Computing Systems. *Computing in science & engineering*, 12(3):66–72, may 2010.
- [11] A.C. Kak and Malcolm Slaney. *Principles of Computerized Tomography*. IEEE Press, 1988.
- [12] Damien Farrell. Radon Transform ImageJ plugin, 2016.
- [13] W. O. Saxton and W. Baumeister. The correlation averaging of a regularly arranged bacterial cell envelope protein. *Journal of Microscopy*, 127(2):127–138, 1982.
- [14] Robert P.J. Nieuwenhuizen, Keith A. Lidke, Mark Bates, Daniela Leyton Puig, David Grünwald, Sjoerd Stallinga, and Bernd Rieger. Measuring image resolution in optical nanoscopy. *Nature Methods*, 10(6):557–562, 2013.
- [15] Nils Gustafsson, Siân Culley, George Ashdown, Dylan M. Owen, Pedro Matos Pereira, and Ricardo Henriques. Fast live-cell conventional fluorophore nanoscopy with ImageJ through super-resolution radial fluctuations. *Nature Communications*, 7:1–9, 2016.
- [16] Eline M Van Dijk, Sule Culha, Mark H Menzen, and Cécile M Bidan. Elastase-Induced Parenchymal Disruption and Airway Hyper Responsiveness in Mouse Precision Cut Lung Slices : Toward an Ex vivo COPD Model. *Frontiers in Physiology*, 7(January):1–11, 2017.
- [17] Alicia Arranz, Di Dong, Shouping Zhu, Charalambos Savakis, Jie Tian, and Jorge Ripoll. In-vivo optical tomography of small scattering specimens: Time-lapse 3D imaging of the head eversion process in *Drosophila melanogaster*. *Scientific Reports*, 4:1–5, 2014.
- [18] Jürgen Mayer, Alexandre Robert-Moreno, Renzo Danuser, Jens V. Stein, James Sharpe, and Jim Swoger. OPTiSPIM: integrating optical projection tomography in light sheet microscopy extends specimen characterization to nonfluorescent contrasts. *Optics letters*, 39(4), 2014.
- [19] J McGinty, H B Taylor, L Chen, L Bugeon, J R Lamb, M J Dallman, and P M French. In vivo fluorescence lifetime optical projection tomography. *Biomed Opt Express*, 2(5):1340–1350, 2011.

- [20] Johnathon R. Walls, Leigh Coultas, Janet Rossant, and R. Mark Henkelman. Three-dimensional analysis of vascular development in the mouse embryo. *PLoS ONE*, 3(8), 2008.
- [21] Jelle van der Horst and Jeroen Kalkman. Image resolution and deconvolution in optical tomography. *Opt. Express*, 24(21):24460–24472, 2016.
- [22] Olli Koskela, Toni Montonen, Birhanu Belay, Edite Figueiras, Sampsa Pursiainen, and Jari Hyttinen. Gaussian Light Model in Brightfield Optical Projection Tomography. *Scientific Reports*, 9(1):1–14, 2019.
- [23] Samuel P.X. Davis, Sunil Kumar, Yuriy Alexandrov, Ajay Bhargava, Gabriela da Silva Xavier, Guy A. Rutter, Paul Frankel, Erik Sahai, Seth Flaxman, Paul M.W. French, and James McGinty. Convolutional neural networks for reconstruction of undersampled optical projection tomography data applied to in vivo imaging of zebrafish. *Journal of Biophotonics*, (August):1–10, 2019.
- [24] Jorge Amich, Zeinab Mokhtari, Marlene Strobel, Elena Vialetto, Dalia Sheta, Yidong Yu, Julia Hartweg, Natarajaswamy Kalleda, Katja J. Jarick, Christian Brede, Ana-Laura Jordán-Garrote, Sina Thusek, Katharina Schmiedgen, Berkan Arslan, Jürgen Pinnecker, Christopher R. Thornton, Matthias Gunzer, Sven Krappmann, Hermann Einsele, Katrin G. Heinze, and Andreas Beilhack. Three-dimensional light sheet fluorescence microscopy of lungs to dissect local host immune-aspergillus fumigatus interactions. *mBio*, 11(1), 2020.
- [25] Isabelle Salwig, Birgit Spitznagel, Marion Wiesnet, and Thomas Braun. Imaging lung regeneration by light sheet microscopy. *Histochemistry and Cell Biology*, (0123456789), 2020.
- [26] Ali Ertürk, Klaus Becker, Nina Jährling, Christoph P. Mauch, Caroline D. Hojer, Jackson G. Egen, Farida Hellal, Frank Bradke, Morgan Sheng, and Hans Ulrich Dodt. Three-dimensional imaging of solvent-cleared organs using 3DISCO. *Nature Protocols*, 7(11):1983–1995, 2012.
- [27] Kwanghun Chung, Jenelle Wallace, Sung Yon Kim, Sandhiya Kalyanasundaram, Aaron S. Andalman, Thomas J. Davidson, Julie J. Mirzabekov, Kelly A. Zalocusky, Joanna Mattis, Aleksandra K. Denisin, Sally Pak, Hannah Bernstein, Charu Ramakrishnan, Logan Grosenick, Viviana Gradinaru, and Karl Deisseroth. Structural and molecular interrogation of intact biological systems. *Nature*, 497(7449):332–337, 2013.

- [28] T. Kuwajima, A. A. Sitko, P. Bhansali, C. Jurgens, W. Guido, and C. Mason. ClearT: a detergent- and solvent-free clearing method for neuronal and non-neuronal tissue. *Development*, 140(6):1364–1368, 2013.
- [29] Meng Tsen Ke, Satoshi Fujimoto, and Takeshi Imai. SeeDB: A simple and morphology-preserving optical clearing agent for neuronal circuit reconstruction. *Nature Neuroscience*, 16(8):1154–1161, 2013.
- [30] Fei Chen, Paul W Tillberg, and Edward S Boyden. Expansion microscopy. *Science*, 347(6221):543 – 548, 2015.
- [31] Sven Truckenbrodt, Christoph Sommer, Silvio O Rizzoli, and Johann G Danzl. A practical guide to optimization in X10 expansion microscopy. *Nature Protocols*, 14(3):832–863, 2019.
- [32] Bo-Jui Chang, Victor Didier Perez Meza, and Ernst H. K. Stelzer. csiLSFM combines light-sheet fluorescence microscopy and coherent structured illumination for a lateral resolution below 100 nm. *Proceedings of the National Academy of Sciences*, 2017.

# Conclusions

The objective of the work presented in this thesis was to develop optical imaging and analysis methods to answer biological questions in models of disease. Questions in three disease models were explored, with biological samples ranging from sub-cellular protein clusters (Part I) to adult mouse lungs (Part II). This variation in length scales, optical properties, resolution, and sensitivity requirements presented unique challenges for each project, requiring tailored combinations of state-of-the-art imaging techniques and analysis methods.

At the nanoscale, the aim was to quantify the proximity between nanoscale distributions of proteins in Parkinson's Disease and HIV *in vitro* models. In Chapter 3, *d*STORM was used successfully to measure nanometre synaptic protein distributions, whereas in Chapter 4 the labelling for *d*STORM was found to be highly unspecific, and instead conventional TIRF microscopy was used to examine viral and ESCRT protein clusters at the membrane. Novel image analysis software tools were developed for both applications which allowed us to remove noise and background signals, select regions of interest, and quantify colocalisation of two labelled species.

At the mesoscale, the goal was to visualise disease phenotypes in large organ samples in 3D. Optical projection tomography was introduced as a bridging technique between medical tomography and conventional 3D light microscopy techniques, as OPT can image both unlabelled anatomical features and fluorescently labelled molecular species in mesoscopic samples. An OPT prototype built by a Masters student cohort was optimised, characterised, and applied to imaging adult mouse lungs in a COPD model.

In summary, the main results from this thesis are:

## **Development of SynaptoAnalysis, a software tool for the detection and analysis of synaptosomes from SMLM data.**

SynaptoAnalysis is a software package developed to measure the clustering of a-syn, and its colocalisation with internalised synaptic vesicles in synaptosomes. SynaptoAnalysis, written in Matlab, contains routines to automatically identify synaptosomes from background in

SMLM data, correct chromatic offset, filter out false positives, and measure colocalisation and cluster size with multiple methods. A model synaptosome was created to generate synthetic SMLM data, which were used to characterise and validate SynaptoAnalysis. Synthetic data sets were useful to benchmark the performance of Ripley's H-function and the RMSD from centroid to measure cluster size.

#### **A-syn colocalises with non-canonically trafficked vesicles.**

A-syn's colocalisation with internalised vesicles was shown to be temperature-independent, which suggests it could be associated with vesicle trafficking mechanisms more likely to be observed at low temperatures when membranes are rigid [1], such as kiss and run. In comparison, VAMP-2's colocalisation with internalised vesicles decreased significantly at sub-physiological temperatures, reinforcing its well established association with SNARE/CM exo/endocytosis. These results provide new insights into the role of a-syn at the pre-synapse, and invite further experiments to determine the identity of the recycling mechanisms in which a-syn could be involved.

#### **A-syn and VAMP-2 clusters increase in size in the absence of calcium.**

Two cluster analysis methods, the RMSD from centroid and Ripley's H-function, showed that a-syn and VAMP-2 clusters imaged by *d*STORM become larger in calcium-depleted levels at the synapse, compared to physiological calcium levels. This is a novel result which could support the hypothesis that a-syn and VAMP-2 interact to tether vesicles together [2].

#### **TIRF microscopy showed EAP45 colocalisation with Gag is dependent on an intact N-terminus.**

SNAP-tag labelled EAP45 and GFP-fused Gag clusters were imaged using TIRF at the cell membrane, and a workflow for performing a nearest-neighbour colocalisation analysis was developed. EAP45 was shown to colocalise with Gag similar to the ALIX protein, which is known to be recruited by Gag during viral egress. Upon deletion of EAP45's N terminus, this colocalisation was significantly decreased. This effect was confirmed using two different cell lines, with similar results. These results provide the first visual evidence of the spatial association between HIV and the ESCRT-II complex, and strengthen previous biochemical evidence [3, 4] showing ESCRT-II is necessary for efficient viral budding.

**A method for visualising co-moving particles revealed three potential classes of dynamic interaction between EAP45 and Gag.**

A novel algorithm for visualising the relative motion of two particles was developed, which allowed us to observe three movement classes of EAP45 clusters relative to Gag clusters: (1) EAP45 oscillates in and out of proximity to Gag; (2) EAP45 comes in close proximity to Gag transiently before moving away; (3) EAP45 remains in close proximity to Gag during the entire observation time. EAP45 particles were found to dwell within ~100 nm of Gag between 10 seconds and 6 mins for all particles analysed. This is the first reported observation of the dynamic interaction between Gag and EAP45; future biophysical studies are necessary to further understand the mechanistic role of ESCRT-II in viral egress.

**Development of a low-cost, open-source OPT system to image large organ samples.**

A low-cost, open-source OPT system, OptiJ, was developed and characterised for the study of disease phenotypes in large organs. 3D reconstructions of excised and cleared mouse lung lobes from a COPD study show high order airways can be visualised within the lung at 50  $\mu\text{m}$  resolution. OptiJ was built using off-the-shelf components, and it uses freely available ImageJ plugins for data pre-processing and reconstruction. A repository containing the raw and reconstructed lung data, instructions for building the instrument, and a detailed software manual was made available to promote the widespread adoption of OPT for imaging large samples.

## References

- [1] Natali L. Chanaday and Ege T. Kavalali. Time course and temperature dependence of synaptic vesicle endocytosis. *FEBS Letters*, 592(21):3606–3614, 2018.
- [2] Jichao Sun, Lina Wang, Huan Bao, Sanjay Premi, Utpal Das, Edwin R. Chapman, and Subhojit Roy. Functional cooperation of  $\alpha$ -synuclein and VAMP2 in synaptic vesicle recycling. *Proceedings of the National Academy of Sciences*, 116(23):201903049, 2019.
- [3] Bo Meng, Natasha C.Y. Ip, Liam J. Prestwood, Truus E.M. Abbink, and Andrew M.L. Lever. Evidence that the endosomal sorting complex required for transport-II (ESCRT-II) is required for efficient human immunodeficiency virus-1 (HIV-1) production. *Retrovirology*, 12(1):1–15, 2015.

- [4] Bo Meng, Natasha C.Y. Ip, Truus E.M. Abbink, Julia C. Kenyon, and Andrew M.L. Lever. ESCRT-II functions by linking to ESCRT-I in human immunodeficiency virus-1 budding. *Cellular Microbiology*, (June 2019):1–15, 2020.



# List of publications

The following preprints and peer-reviewed publications contain work described in this thesis:

- Meng, B., **Vallejo Ramirez, P. P.**, Scherer, K. M., Bruggeman, E., Kenyon, J. C., Kaminski, C. F., & Lever, A. M. (2020). *Distinct domain requirements for EAP45 in HIV budding, late endosomal recruitment, and cytokinesis*. bioRxiv, 2020.05.23.112607. <https://doi.org/10.1101/2020.05.23.112607> (submitted for peer-review)
- **Vallejo Ramirez, P.P.**, Zammit, J., Vanderpoorten, O. et al. (2019). *OptiJ: Open-source optical projection tomography of large organ samples*. Scientific Reports, 9, 15693. <https://doi.org/10.1038/s41598-019-52065-0>
- Rowlands, C. J., Ströhl, F., **Ramirez, P. P. V.**, Scherer, K. M., & Kaminski, C. F. (2018). *Flat-Field Super-Resolution Localization Microscopy with a Low-Cost Refractive Beam-Shaping Element*. Scientific Reports, 8(1), 5630. <https://doi.org/10.1038/s41598-018-24052-4>

Other preprints and publications containing work not included in this thesis:

- Mela, I., **Vallejo-Ramirez, P.P.**, Makarchuk, S., Christie, G., Bailey, D., Henderson, R.M., Sugiyama, H., Endo, M. and Kaminski, C.F. (2020). *DNA Nanostructures for Targeted Antimicrobial Delivery*. Angew. Chem. Int. Ed.. <https://doi:10.1002/anie.202002740>
- Hockings, C., Poudel, C., Feeney, K.A., Novo, C.L., Hamouda, M.S., Mela, I., Fernandez-Antoran, D., **Vallejo-Ramirez, P.P.**, Rugg-Gunn, P.J., Chalut, K., Kaminski, C.F., Kaminski-Schierle, G. (2020). *Illuminating Chromatin Compaction in Live Cells and Fixed Tissues Using SiR-DNA Fluorescence Lifetime*. bioRxiv: 2020.05.02.073536. <http://biorxiv.org/content/early/2020/05/02/2020.05.02.073536.abstract>
- Koppers, M., Cagnetta, R., Shigeoka, T., Wunderlich, L. C., **Vallejo-Ramirez, P.**, Qiaojin Lin, J., ... Holt, C. E. (2019). *Receptor-specific interactome as a hub for*

*rapid cue-induced selective translation in axons*. eLife, 8, 1–27. <https://doi.org/10.7554/eLife.48718>

- Dolan, S., Kohlstedt, M., Trigg, S., **Ramirez, P. V.**, Wittmann, C., Kaminski, C., & Welch, M. (2019). *Contextual flexibility in Pseudomonas aeruginosa central carbon metabolism during growth in single carbon sources*. mBio, (March), 1–20. <https://doi.org/10.1101/828012>